Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

Инженерно-экономический факультет

Кафедра экономической информатики

ЭКОНОМЕТРИКА И ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ И МОДЕЛИ

В двух частях
Часть 1

О. А. Голда, Н. М. Матвейчук

ЭКОНОМЕТРИКА

*Рекомендовано УМО по образованию
в области информатики и радиоэлектроники
в качестве учебно-методического пособия
для практических занятий
для специальности 1-28 01 01 «Экономика электронного бизнеса»*

Минск БГУИР 2019

УДК 519.862.6(075)
ББК 65в6я7
Э40

Рецензенты:

кафедра экономической информатики учреждения образования
«Белорусский государственный экономический университет»
(протокол №4 от 25.10.18);

доцент кафедры маркетинга
Белорусского национального технического университета
кандидат технических наук, доцент Б. А. Железко

Э40 **Эконометрика** и экономико-математические методы и модели. В 2 ч. Ч. 1 : Эконометрика : учеб.-метод. пособие / О. А. Голда, Н. М. Матвейчук. – Минск : БГУИР, 2019. – 154 с. : ил.
ISBN 978-985-543-474-1 (ч. 1).

Излагается теоретический материал и приводятся примеры построения эконометрических моделей, а также задания для самостоятельного решения по таким разделам эконометрики, как линейные регрессионные модели, эконометрический анализ в условиях нарушений предпосылок МНК, решение систем линейных уравнений.

УДК 519.862.6(075)
ББК 65в6я7

ISBN 978-985-543-474-1 (ч. 1)
ISBN 978-985-543-473-4

Содержание

Практическая работа №1
Статистическая обработка результатов наблюдений

1.1 Основные понятия и определения

Пусть по известным способам отбора статистических данных создан массив экспериментальных наблюдений. Однако в том виде, в каком эти данные представлены, они мало пригодны для дальнейшей обработки. Гораздо больших результатов можно достичь, представляя эти данные в виде ряда распределения.

Статистической совокупностью (или признаком – это свойство элементов совокупности) называется масса единиц, обладающих некоторыми общими свойствами, существенными для их характеристики.

Значения признака у отдельных членов совокупности называются вариантами.

Операция упорядочения статистической совокупности, т. е. расположение вариантов в порядке возрастания (или убывания), называется ранжированным рядом.

Числа, которые характеризуют количество повторений каждого значения признака в данной совокупности, называются абсолютными частотами признака и обозначаются n_i. Сумма $n = \sum n_i$ частот называется объемом совокупности.

Отношение $w_i = \dfrac{n_i}{n}$ соответствующих частот к объему совокупности называется относительными частотами (или частостями).

Вариационным рядом называется упорядоченный ряд различных значений варьирующего признака и соответствующих им частот (или частостей).

Вариационные ряды могут быть дискретными и непрерывными. Прежде чем приступить к обработке статистического материала, его подвергают контролю, группировке и разбиению на интервалы.

Под группировкой понимают не просто расчленение массы элементов на группы, а такое, при котором группы будут состоять из качественно однородных элементов. Так, если число групп (или интервалов) обозначить через K, то для элементов совокупности $n \le 100$ число интервалов определяется по формуле $K = 1 + 3{,}322\lg n$, а для $n > 100$ – по формуле $K \le 5\lg n$.

В непрерывном вариационном ряде возможные значения задаются интервалами «от … до», и в таких рядах частоты относятся не к какому-нибудь отдельному значению признака, а ко всему интервалу.

Разности $h_i = \alpha_2 - \alpha_1$, …, $h_m = \alpha_{m+1} - \alpha_m$, где $\alpha_1, \alpha_2, ..., \alpha_{m+1}$ – граничные точки интервалов, называются интервальными разностями или шагом. Наиболее простыми являются вариационные ряды с равными интервалами, т. е. для них $h = \text{const}$.

Частота, приходящаяся на единицу величины интервала варьирования признака, называется абсолютной плотностью частоты и обозначается $П_а = \frac{n_i}{h_i}$. Отношение частостей к соответствующим интервальным разностям называется плотностью относительной частоты и обозначается $П_о = \frac{w_i}{h_i}$.

Накопленными частотами (или частостями) называется сумма частот (или частостей) всех предшествующих интервалов, включая данный, они обозначаются N_{ni} или N_{wi}.

Заметим, что в качестве варианта непрерывного вариационного ряда берут середины соответствующих интервалов, т. е. $X_i^* = \frac{\alpha_i + \alpha_{i+1}}{2}$.

1.2 Порядок выполнения работы

1 Для построения непрерывных вариационных рядов нужно установить полную шкалу интервалов, т. е. найти k и h. Для определения оптимальной величины шага интервала, т. е. такой, при которой построенный интервальный ряд не был бы слишком громоздким и в то же время позволял выявить характерные черты рассматриваемого явления, можно использовать формулу Стэрджеса:

$$h = \frac{X_{\max} - X_{\min}}{1 + 3{,}322 \cdot \lg n} = \frac{R}{K}, \tag{1.1}$$

где $X_{\max}$ и $X_{\min}$ – соответственно максимальная и минимальная варианты;

$R = X_{\max} - X_{\min}$ – размах вариации, который представляет собой амплитуду колебания или широту рассеяния и является величиной неустойчивой, зависящей от случайных обстоятельств;

K – число интервалов при объеме выборки $n \le 100$.

Если h оказывается дробным числом, то за единицу интервала следует взять ближайшее целое число. За начало первого интервала рекомендуется принимать величину, равную $X_{\min} - \frac{h}{2}$, т. е. $a_1 = a_1 + h$ и т. д. до тех пор, пока начало следующего по порядку интервала не будет равным или большим $X_{\max}$. Для подсчета частот вариантов в интервал включаются варианты, большие чем нижняя граница (а для первого интервала включают и a_1) и меньшие или равные верхней границе интервала, т. е. верхнюю границу включают в интервал.

2 Все расчетные таблицы можно составлять произвольно (самостоятельно). Так, после проведения расчетов, выполненных в пункте 1, можно рекомендовать следующую расчетную таблицу 1.1.

Таблица 1.1 – Расчет частот

Граница интервалов	Частота n_i	Частота w_i	Середины интервалов X_i^*	Накопленные частоты (или частости) $N_{ni}(N_{wi})$

3 Для графического изображения вариационных рядов применяются: полигон частот (или частостей) и гистограмма плотности частот (или относительных частот).

Чтобы построить полигон в прямоугольной системе координат, нужно отложить точки $M_i\left[x_i^*, n_i(\text{или } w_i)\right]$, а затем соединить их последовательно прямолинейными отрезками. Полученная ломаная линия и называется полигоном частот (или частостей).

Чаще всего полигон применяется для изображения дискретного вариационного ряда, но может быть применен и для интервального ряда. Сумма длин всех ординат полигона равна численности вариационного ряда *n*. Полигон может быть замкнутым.

Для построения гистограммы в системе *XOY* нужно по оси абсцисс отложить отрезки, изображающие интервалы значений варьирующего признака. На этих отрезках, как на основаниях, построить прямоугольники, высоты которых пропорциональны частотам: $W_i = \Pi_\text{а} = \frac{n_i}{n}$, где $\Pi_\text{а}$ – плотность частоты, или частости:

$$\Pi_\text{а} = \frac{W_i}{n}. \tag{1.2}$$

В результате получится ступенчатая фигура в виде сдвинутых друг к другу прямоугольников, площади которых пропорциональны частотам (или частостям).

4 Чтобы определить параметры $\bar{x}_{\text{В}}$, $D_{\text{В}}$, $\sigma_{\text{В}}$, нужно обратиться к методу произведения (или методу условного нуля), для чего составляют расчетную таблицу. Эту таблицу составляют следующим образом:

- в первый столбец записывают выборочные варианты x_i*, располагая их в возрастающем порядке;

- во второй столбец записывают частоты вариант. Складывают эти частоты и их сумму (объем выборки n) помещают в нижнюю клетку столбца;

- в третий столбец записывают условные варианты u_i. В качестве условного (ложного) нуля C выбирают варианту с наибольшей частотой. Практически третий столбец может быть заполнен так: в клетке строки, содержащей наибольшую частоту, записать 0; в клетках над нулем записать последовательно -1, -2, -3 и т. д., а под нулем записать 1, 2, 3, и т. д. Заметим, что в качестве условного нуля выгодно принимать варианту, которая примерно находится в середине столбца (даже если эта варианта и не имеет наибольшей частоты);

- умножить частоты на условные варианты и записать их произведения n_i u_i в четвертый столбец таблицы. Сложив все полученные числа, их сумму $\sum n_i u_i$ поместить в нижнюю клетку столбца;

- умножить частоты на квадраты условных вариант и записать их произведения $n_i u_i^2$ в пятый столбец. Сложив все полученные числа, их сумму $\sum n_i u_i^2$ поместить в нижнюю клетку столбца;

- умножить частоты на квадраты условных вариант, каждая из которых увеличена на единицу и записать произведения $n_i(u_i+1)^2$ в шестой столбец. Сложив все полученные числа, их сумму $\sum n_i(u_i+1)^2$ поместить в нижнюю клетку столбца;

- вычислить условные моменты первого и второго порядков и записать их в нижних клетках седьмого столбца.

Контроль: если сумма $\sum n_i(u_i+1)^2$ окажется равной сумме $\sum n_i u_i^2 + 2\sum n_i u_i + n$, то вычисления проведены правильно.

В итоге получим расчетную таблицу 1.2.

Таблица 1.2 – Расчет условных моментов

x_i^*	n_i	u_i	$n_i \cdot u_i$	$n_i \cdot u_i^2$	$n_i(u_i+1)^2$	Условные моменты	
						M_1^*	M_2^*

Для расчета параметров используются следующие формулы:

$\bar{x}_{\text{в}} = M_1^* \cdot h + C$ – выборочная средняя;

$D_{\text{в}} = \left[M_2^* - \left(M_1^* \right)^2 \right] \cdot h^2$ – выборочная дисперсия;

h – шаг; С – ложный нуль;

$u_i = \dfrac{x_i^* - C}{h}$ – условная варианта;

$M_1^* = \dfrac{\sum n_i u_i}{n}$ – условный момент 1-го порядка;

$M_2^* = \dfrac{\sum n_i u_i^2}{n}$ – условный момент 2-го порядка;

$\sigma_{\text{в}} = \sqrt{D_{\text{в}}}$ – выборочное среднее квадратическое отклонение;

$V_{\sigma_{\text{В}}} = \dfrac{\sigma_{\text{в}}}{\bar{x}_{\text{в}}} \cdot 100\ \%$ – коэффициент вариации по среднему квадратическому отклонению – это относительная величина, служащая для характеристики колебания признака, представляет собой процентное отношение абсолютных величин: среднего квадратического отклонения к средней арифметической.

5 Доверительный интервал для математического ожидания с надежностью $\gamma = 0{,}95$ (при нормальном распределении) можно определить так:

$$\bar{x}_{\text{в}} - t \cdot \frac{\sigma_{\text{в}}}{\sqrt{n}} < a < \bar{x}_{\text{в}} + t \cdot \frac{\sigma_{\text{в}}}{\sqrt{n}}. \tag{1.3}$$

Параметр t находим из равенств

$$P \cdot \left(\bar{x}_{\text{в}} - t \cdot \frac{\sigma_{\text{в}}}{\sqrt{n}} < a < \bar{x}_{\text{в}} + t \cdot \frac{\sigma_{\text{в}}}{\sqrt{n}} \right) = 2 \cdot \Phi(t), \tag{1.4}$$

$$2\Phi(t) = \gamma \quad \text{или} \quad \Phi(t) = \frac{\gamma}{2}.$$

Значение t находится по таблице значений функции Лапласа $\Phi(x)$, которому соответствует значение функции, равное $\dfrac{\gamma}{2}$, т. е. при $\gamma = 0{,}95$; $\dfrac{\gamma}{2} = 0{,}475$ число $t = 1{,}96$.

Для определения трехсигмового интервала $\left(\bar{x}_{\text{в}} - 3\sigma_{\text{в}}, \bar{x}_{\text{в}} + 3\sigma_{\text{в}} \right)$ можно воспользоваться неравенствами

$$\bar{x}_{\text{в}} - 3\sigma_{\text{в}} < a < \bar{x}_{\text{в}} + 3\sigma_{\text{в}}. \tag{1.5}$$

Для проверки гипотезы о нормальном распределении статистической совокупности при уровне значимости $\alpha = 0{,}05$ воспользуемся критерием χ^2 Пирсона. Рассмотрим критическую область:

$$P\left[\chi^2 > \chi^2_{\text{кр}}(\alpha, k) \right] = \alpha. \tag{1.6}$$

Вычисляем теоретические частоты:

$$n_i' = \frac{n \cdot h}{\sigma_{\text{В}}} \cdot \varphi(u_i),$$

где $u_i = \dfrac{x_i^* - \bar{x}_{\text{В}}}{\sigma_{\text{В}}}$;

$\varphi(u_i) = \dfrac{1}{\sqrt{2\pi}} \cdot e^{-\frac{u_i^2}{2}}$ – функция Гаусса, значения которой находятся по специальной таблице.

Для сравнения эмпирических и теоретических частот с помощью критерия χ^2 Пирсона составляем следующие расчетные таблицы 1.3 и 1.4.

Таблица 1.3 – Расчет теоретических частот

i	x_i^*	u_i	$\varphi(u_i)$	$n_i' = \frac{n \cdot h}{\sigma_{\text{В}}} \cdot \phi(u_i)$

Таблица 1.4 – Предварительные расчеты для вычисления χ^2

i	n_i	n_i'	$n_i - n_i'$	$(n_i - n_i')^2$	$\frac{(n_i - n_i')^2}{n_i'}$	n_i^2	$\frac{n_i^2}{n_i'}$

По таблице 1.4 вычисляем

$$\chi^2_{\text{набл}} = \sum_{i=1}^{n} \frac{(n_i - n_i')^2}{n_i'} . \tag{1.7}$$

Из таблицы А.4 (приложение А) критических точек распределения χ^2 по заданному уровню значимости α и числу степеней свободы $k = s - 1 - r$, где s – число групп, r – число параметров выбранного распределения, находим критическую точку $\chi^2_{\text{кр}}(\alpha , k)$ правосторонней или левосторонней критической области.

Если $\chi^2_{\text{набл}} \le \chi^2_{\text{кр}}$, то нет оснований отвергать гипотезу о нормальном распределении признака.

Если $\chi^2_{\text{набл}} > \chi^2_{\text{кр}}$, то гипотезу о нормальном распределении признака отвергают.

Для контроля вычислений осуществляют проверку

$$\sum_{i=1}^{k} \frac{\left(n_i - n_i'\right)}{n_i'} = \sum_{i=1}^{k} \frac{n_i^2}{n_i'} - n. \tag{1.8}$$

Для графической иллюстрации полученных результатов нужно найти асимметрию и эксцесс.

Распределение считается почти симметричным, если $|A_s| < 0{,}1$, и сильно асимметричным, если $|A_s| > 0{,}5$, где A_s – асимметрия. Если $A_s < 0$, то длинная часть графика и $M(x) = a$ расположена левее $\text{Mo}X$, а крутая часть графика – справа. Здесь $\text{Mo}X$ – мода признака.

Если $A_s > 0$, то длинная часть кривой и $M(x)$ расположены правее моды, а крутая часть – слева.

Асимметрия вычисляется следующим образом:

$$A_3 = \frac{M_3}{\sigma_{\text{В}}^3}, \tag{1.9}$$

где $M_3 = m_3 = \dfrac{\sum n_i \left(x_i - \bar{x}_B\right)^3}{n}$ – центральный момент 3-го порядка;

$M_4 = m_4 = \dfrac{\sum n_i \left(x_i - \bar{x}_B\right)^4}{n}$ – центральный момент 4-го порядка.

С равным шагом h эти моменты удобно вычислять по формулам:

$$M_3 = \left[M_3^* - 3M_1^* M_2^* + 2\left(M_1^*\right)^3 \right] \cdot h^3, \tag{1.10}$$

$$M_4 = \left[M_4^* - 4M_1^* M_3^* + 6\left(M_1^*\right)^2 \cdot M_2^* - 3\left(M_1^*\right)^4 \right] \cdot h^4, \tag{1.11}$$

где $M_k^* = \dfrac{\sum n_i u_i^k}{n}$, $u_i = \dfrac{x_i - c}{h}$.

Эксцесс определяется с помощью равенства $E_k = \dfrac{M_4}{\sigma_{\text{В}}^4} - 3$.

Если $E_k > 0$, то кривая имеет более высокую и более «острую» вершину, чем стандартная нормальная кривая.

Если $E_k < 0$, то кривая имеет более низкую и более «плоскую» вершину по сравнению со стандартной нормальной кривой.

С учетом A_s и E_k возможны следующие случаи:

1) $A_s > 0$, $E_k > 0$; 2) $A_s > 0$, $E_k < 0$;

3) $A_s < 0$, $E_k > 0$; 4) $A_s < 0$, $E_k < 0$.

В заключение нужно сделать вывод о принятии или непринятии выдвинутой гипотезы.

1.3 Варианты заданий

По данным статистической выборки:

1 Составить:

а) вариационный ряд;

б) число интервалов и шаг.

2 При выполнении вычислений (там, где это необходимо) составить расчетные таблицы.

3 Построить гистограмму плотности частоты (или частостей).

4 Определить методом произведений (методом условного нуля) следующие параметры:

а) среднюю выборочную $\bar{x}_{\text{в}}$;

б) выборочную дисперсию $D_{\text{в}}$;

в) выборочное среднее квадратическое отклонение $\sigma_{\text{в}}$;

г) коэффициент вариации $V_{\sigma_{\text{в}}}$ по среднему квадратическому отклонению.

5 Считая распределение нормальным (гипотеза), определить:

а) доверительный интервал для математического ожидания с надежностью $\gamma = 0{,}95$;

б) трехсигмовый интервал.

6 Проверить при заданном уровне значимости $\alpha = 0{,}05$ с помощью критерия χ^2 Пирсона, что данный признак распределен по нормальному закону.

7 Найти асимметрию, эксцесс и построить график.

8 Сделать вывод о принятии или непринятии выдвинутой гипотезы о нормальном распределении выборки.

Вариант 1

24,04	29,66	33,00	24,27	40,05	25,89	29,48	31,37	24,86	28,68
28,06	26,84	29,15	25,06	31,35	24,68	27,66	21,66	19,92	22,17
35,28	25,68	28,71	30,71	29,27	24,08	30,76	30,73	29,06	29,86
24,34	20,09	32,21	30,24	31,99	40,19	35,67	25,93	28,58	28,94
35,16	33,68	27,95	33,91	33,06	40,33	26,10	32,54	30,94	23,00
23,11	34,90	21,91	27,80	19,94	22,89	34,20	31,71	35,14	29,16
32,99	30,45	29,07	25,14	32,35	36,55	26,60	25,64	32,99	26,67
29,10	32,67	34,71	33,59	35,26	36,49	26,57	32,29	24,65	37,92
37,46	25,08	21,58	32,62	33,81	20,93	27,40	29,75	31,46	37,56
20,83	24,16	12,76	30,33	31,39	17,38	33,69	21,25	34,15	25,57

Вариант 2

33,86	32,37	36,26	26,36	37,58	41,47	31,95	37,65	28,00	29,12
35,20	23,71	25,83	27,80	30,22	24,94	32,10	35,16	33,65	18,43
29,45	35,61	20,94	34,89	35,15	26,87	23,52	30,34	27,52	23,39
29,59	31,83	33,57	35,14	22,90	31,13	33,96	25,18	33,24	37,68
24,78	24,78	23,78	30,90	26,50	27,05	28,71	31,00	33,01	38,80
28,98	27,11	30,91	21,89	30,83	28,33	27,44	27,46	18,51	27,99
30,93	36,27	31,90	24,80	26,93	31,40	24,83	31,67	31,48	23,09
28,43	24,57	31,74	30,40	28,33	27,98	29,20	30,06	23,93	35,40
33,79	33,54	31,96	23,76	31,46	32,02	31,64	22,43	35,66	26,70
26,80	29,93	37,51	19,04	28,43	31,72	25,25	34,68	27,82	33,62

Вариант 3

32,00	23,38	34,32	31,86	28,60	26,08	33,06	43,98	36,26	25,44
29,98	30,87	32,04	26,36	21,69	36,14	35,55	37,99	38,45	27,81
23,03	23,27	32,65	34,30	35,44	26,17	23,62	25,06	27,67	31,59
29,40	26,46	20,41	24,34	29,75	31,55	31,45	30,78	36,63	22,48
40,44	26,78	34,98	33,83	30,11	30,24	36,48	24,98	25,94	27,91
32,41	31,51	32,25	25,53	34,65	33,17	36,65	22,55	31,37	35,64
26,23	31,94	17,37	33,93	28,81	32,47	29,41	39,36	28,96	37,69
32,88	20,09	30,52	29,50	30,83	27,91	24,98	30,29	37,13	29,25
20,92	22,13	34,06	29,15	17,84	32,18	25,37	23,41	27,27	22,17
37,68	34,51	29,50	29,75	27,86	35,39	37,83	28,36	27,29	27,25

Вариант 4

21,49	32,69	37,77	30,91	21,79	34,38	32,26	37,63	22,13	30,03
34,29	29,33	32,65	19,02	25,13	39,80	34,23	25,80	35,86	30,25
26,96	31,17	31,91	26,89	25,58	23,83	37,46	32,16	27,75	29,44
29,11	36,49	35,75	25,56	35,16	31,51	19,00	39,70	28,68	24,49
27,80	31,83	24,13	28,68	19,16	30,50	33,94	26,51	32,71	29,20
30,01	39,51	28,35	25,57	27,20	45,52	31,96	27,98	29,59	28,32
31,21	29,35	30,09	33,15	32,73	27,59	28,99	24,82	36,60	20,01
36,98	24,36	35,93	33,18	27,43	33,46	25,78	34,64	21,64	29,05
28,18	37,97	33,34	30,85	26,63	27,10	31,42	35,72	30,58	26,95
30,42	35,45	29,98	32,92	37,87	29,02	31,73	33,68	44,75	26,97

Вариант 5

27,59	31,56	30,97	27,07	32,57	35,12	23,13	30,03	23,14	28,93
27,66	22,34	33,57	33,11	37,19	34,71	25,54	36,98	27,51	37,86
34,99	31,47	27,38	29,01	28,48	37,97	24,07	30,32	28,14	34,85
34,84	23,97	26,10	23,28	28,63	30,90	34,38	31,69	21,16	29,58
28,81	36,12	19,55	21,83	20,50	29,67	27,28	36,51	30,30	33,95
40,47	33,47	20,32	26,60	30,85	24,47	33,05	36,55	39,15	33,83
26,07	27,62	26,15	30,60	30,73	19,20	31,35	30,27	28,89	33,89
23,11	33,60	32,80	31,46	27,04	21,94	28,30	26,48	24,07	39,26
19,82	28,95	25,36	21,08	34,60	28,62	29,52	36,56	30,88	31,36
27,30	29,44	32,51	30,18	28,28	32,28	28,11	33,55	23,26	28,39

Вариант 6

31,00	22,94	23,47	28,48	31,33	27,18	18,97	25,73	38,35	29,64
24,27	25,14	19,42	32,13	30,95	27,17	26,02	26,11	34,17	25,42
33,27	24,53	34,89	28,04	29,83	28,48	36,66	25,73	33,54	34,29
20,54	28,04	26,01	24,54	30,84	19,42	31,10	27,99	39,31	28,30
29,34	29,52	33,14	31,04	25,93	23,57	25,05	34,40	28,75	30,47
22,50	29,97	36,71	21,47	27,38	28,05	34,03	28,92	33,14	25,43
33,65	24,55	35,80	23,06	33,70	37,23	26,86	34,75	35,28	26,95
30,71	25,01	38,40	26,50	19,13	31,81	25,95	25,21	26,51	40,71
28,98	27,02	26,30	25,53	39,05	21,95	26,90	26,16	31,04	25,90
27,00	28,83	30,05	40,11	27,05	26,57	37,22	28,94	28,35	29,44

Вариант 7

29,56	32,80	27,27	35,49	40,23	34,70	33,49	20,98	30,49	27,46
30,32	34,77	39,68	30,68	31,63	37,75	19,61	28,76	32,47	35,11
21,21	36,18	21,26	27,32	33,53	34,40	22,96	29,95	27,70	31,91
34,31	27,58	33,09	31,41	31,29	26,10	29,22	32,40	24,12	25,51
32,33	36,02	31,24	35,52	33,89	24,19	23,90	22,17	34,06	35,58
22,31	30,23	14,82	27,21	36,94	34,35	27,09	31,31	24,58	26,77
32,50	34,,93	32,60	39,52	31,55	26,19	31,33	32,81	42,82	33,71
25,58	33,55	29,60	36,74	33,47	20,38	34,00	42,87	24,55	31,00
36,71	37,05	34,77	32,81	32,69	31,39	37,16	35,07	32,50	32,13
16,48	34,86	33,58	19,54	24,05	28,68	18,37	34,53	28,23	19,85

Вариант 8

32,06	34,21	25,18	19,77	31,89	21,24	37,02	34,57	37,94	28,52
21,15	25,80	22,80	37,61	39,37	26,85	36,72	26,94	28,61	31,81
17,54	22,04	34,74	26,11	28,77	32,29	32,00	40,99	29,13	34,83
36,89	26,93	29,28	29,23	30,58	32,73	28,40	27,77	26,99	39,63
30,57	30,36	27,94	39,23	27,45	30,84	31,82	22,81	34,55	30,82
20,23	25,52	22,84	27,93	35,68	24,05	25,71	16,86	26,72	40,35
26,43	22,54	31,02	27,58	17,26	28,82	29,65	25,83	24,49	23,60
30,03	22,61	43,71	25,02	32,68	33,42	30,01	24,78	30,03	35,39
20,30	38,92	33,69	36,08	37,55	29,20	28,76	32,90	33,56	25,39
22,30	33,25	35,81	24,52	30,78	33,88	36,38	25,63	36,29	33,96

Вариант 9

26,66	23,35	26,06	32,20	25,57	27,75	33,19	22,18	22,61	23,59
37,79	34,27	27,63	28,23	37,30	20,97	34,15	34,34	28,57	27,08
27,65	26,34	29,28	29,13	23,00	31,57	27,42	29,58	28,89	30,72
26,09	34,16	26,54	29,84	30,00	27,83	27,68	37,65	23,64	32,23
33,74	22,66	31,53	29,67	37,63	31,88	30,09	27,89	27,63	33,01
32,57	28,72	32,13	27,13	32,11	38,89	28,99	35,17	29,64	28,59
29,28	31,80	29,14	36,45	33,29	22,72	31,07	40,58	23,34	24,70
27,25	25,63	27,72	17,05	37,39	19,70	20,58	33,45	29,26	36,54
32,58	30,43	30,32	31,37	31,21	37,46	35,73	23,87	36,60	32,87
20,88	35,50	24,30	34,05	25,29	30,33	26,03	31,85	35,73	27,04

Вариант 10

31,50	27,31	27,89	27,99	29,75	35,15	25,29	33,09	27,73	37,61
31,72	33,26	30,92	27,96	27,71	29,46	28,73	25,79	29,19	35,25
37,61	26,78	29,75	30,22	37,52	28,79	26,89	34,51	26,26	33,71
33,20	27,59	15,59	27,64	32,81	36,79	30,60	37,05	17,12	32,47
22,15	28,85	40,50	24,76	25,15	32,89	31,20	25,11	31,73	29,83
25,62	19,59	29,61	28,27	27,85	36,80	32,02	33,10	27,90	31,25
28,60	32,04	30,79	31,51	31,08	31,75	24,23	21,74	32,59	37,24
27,60	24,48	28,09	29,79	37,41	33,84	28,08	34,18	32,31	30,87
21,17	29,60	24,99	24,89	28,17	29,65	32,63	28,62	31,64	32,05
29,29	24,27	26,50	28,05	30,06	24,65	31,74	28,42	31,28	30,84

Вариант 11

29,66	41,73	23,65	28,57	29,49	31,71	33,46	23,94	30,54	33,02
37,60	25,75	31,73	30,10	27,48	24,29	37,26	33,56	30,26	20,66
26,54	26,72	20,54	23,15	24,87	30,79	27,59	15,86	20,48	29,17
34,70	31,43	31,59	25,32	29,61	31,03	34,98	30,31	28,42	19,03
32,80	23,97	23,59	27,76	28,74	40,79	32,88	25,56	38,06	34,54
32,73	34,31	32,69	26,01	32,45	32,25	31,80	32,23	27,49	21,14
27,38	28,69	27,08	30,57	32,87	22,55	24,70	33,46	30,81	27,79
28,34	29,68	33,05	30,31	32,34	32,92	25,30	26,61	25,47	24,63
28,62	33,34	28,82	31,50	39,75	33,19	16,61	24,03	33,91	34,23
36,85	29,32	36,98	29,66	30,45	26,05	21,97	24,69	27,53	25,06

Вариант 12

26,72	31,22	32,72	32,12	29,24	23,02	32,73	30,43	33,90	27,74
31,52	33,65	32,90	22,83	25,87	25,46	35,63	26,90	32,26	36,71
35,22	31,73	35,02	32,60	28,02	22,80	32,93	28,80	18,91	20,26
25,38	33,10	31,17	34,10	31,82	26,91	26,34	27,92	28,95	33,67
28,08	30,70	30,02	25,11	33,99	25,06	28,73	33,46	34,54	25,69
25,67	28,20	31,54	27,78	34,97	31,38	32,74	28,34	30,04	29,74
38,18	32,39	26,60	23,48	27,97	32,05	22,67	37,40	29,47	31,12
37,65	28,38	37,95	38,14	35,94	26,87	34,31	35,73	33,76	34,51
20,57	31,04	30,21	37,29	22,91	25,53	22,87	26,50	33,08	29,12
35,06	33,71	20,43	34,65	21,69	26,68	30,07	21,31	26,73	29,47

Вариант 13

29,47	28,87	27,98	30,72	31,82	34,34	29,82	31,90	31,01	32,91
28,38	23,87	18,34	27,86	28,96	33,97	36,93	35,64	30,67	26,37
38,65	32,68	31,24	28,68	30,03	30,69	38,46	31,88	28,23	33,04
21,06	26,10	21,38	29,48	32,17	36,36	26,99	32,20	28,19	35,58
26,34	30,11	26,55	37,59	26,18	30,48	38,04	38,48	29,54	40,72
34,56	23,55	25,58	27,05	35,84	18,85	16,16	32,19	34,36	26,09
27,68	38,05	29,66	22,93	34,41	36,67	29,45	37,19	28,46	33,25
35,76	31,40	23,88	24,12	34,45	33,88	22,48	32,39	31,44	26,63
31,93	31,11	31,05	28,23	26,35	35,10	27,50	25,24	26,89	31,51
31,37	30,99	33,39	31,11	37,15	31,74	24,19	34,55	34,72	31,46

Вариант 14

29,37	28,74	31,79	24,14	26,04	23,42	30,97	26,27	37,59	27,86
27,13	28,40	15,19	28,84	32,56	32,92	31,14	28,89	28,91	27,16
31,88	32,03	30,69	31,44	24,59	35,61	36,20	25,13	28,29	30,21
27,38	30,50	34,22	25,14	31,49	26,42	25,90	29,29	28,33	26,26
35,56	37,46	37,15	32,40	32,90	36,13	34,76	24,51	30,34	14,87
28,52	30,73	29,78	34,63	33,66	35,87	30,23	34,27	34,42	26,03
19,64	34,90	30,90	31,07	28,19	26,34	33,17	37,14	26,84	37,35
46,15	27,43	31,86	28,50	24,83	33,63	33,90	38,06	34,73	23,36
24,39	29,41	32,13	31,16	30,50	36,72	31,85	26,16	27,52	25,93
40,88	31,39	19,95	27,15	30,71	29,63	30,83	28,03	31,98	36,26

Вариант 15

23,96	33,21	34,59	25,14	31,98	34,14	33,65	33,72	24,07	30,04
31,87	37,23	28,10	33,63	25,98	27,08	35,47	27,38	36,77	35,00
33,95	33,40	20,71	23,44	28,55	26,14	32,01	24,00	18,09	32,22
29,32	30,16	30,53	29,29	35,50	27,19	33,04	34,58	28,14	22,93
32,77	31,71	35,07	27,57	30,97	26,96	41,18	27,54	30,66	24,87
29,10	32,81	32,61	26,52	27,46	30,24	26,74	30,93	25,97	25,81
31,35	45,79	19,40	25,37	36,16	26,79	31,78	37,96	26,21	25,63
27,47	34,68	23,88	34,26	23,46	30,97	20,33	30,63	27,55	35,22
22,56	27,41	31,57	30,66	37,34	33,95	26,42	32,64	29,25	32,84
33,51	35,64	30,38	32,00	24,89	22,31	33,80	22,36	33,71	29,20

Вариант 16

35,86	14,89	29,07	28,34	20,54	26,23	28,93	25,36	23,02	36,17
26,05	37,85	25,40	26,12	29,07	28,56	35,70	26,12	26,87	31,41
23,92	30,01	32,80	24,77	32,79	27,99	30,50	36,39	30,33	30,43
23,74	28,38	33,09	25,51	13,56	33,98	33,13	26,78	30,16	33,54
30,03	29,37	28,84	27,83	35,59	29,77	28,58	38,63	38,69	37,46
35,58	34,01	33,55	29,22	33,65	38,15	24,56	22,83	25,06	23,05
29,45	35,59	34,63	28,56	29,55	30,68	29,70	24,58	29,01	42,18
28,72	40,36	20,11	35,62	32,15	25,20	19,71	28,82	32,92	32,42
27,73	35,73	31,94	24,76	19,41	33,00	34,70	28,71	23,54	24,56
32,43	20,95	36,58	21,63	26,00	26,47	32,66	35,09	29,84	21,33

Вариант 17

32,09	27,43	34,37	27,39	36,65	18,57	33,68	34,86	29,06	26,89
23,97	32,92	36,07	22,71	36,57	22,31	26,90	30,05	26,01	36,91
27,23	37,59	31,07	29,31	33,23	34,71	31,85	26,23	31,07	29,30
30,21	28,94	31,86	47,65	28,87	36,40	30,79	30,30	27,73	25,99
22,27	24,53	41,39	23,39	28,23	37,83	38,21	33,76	28,85	35,19
26,82	30,89	31,01	29,07	30,83	25,69	28,91	35,68	21,84	30,34
34,06	21,94	41,18	31,83	37,67	31,63	33,95	38,56	23,34	18,95
28,37	26,08	25,63	38,94	32,31	29,18	37,29	25,01	24,40	33,41
32,53	34,45	34,06	40,63	25,02	31,09	29,84	25,36	27,27	22,42
21,24	23,53	32,29	27,20	30,09	36,83	22,26	36,45	36,67	33,87

Вариант 18

27,93	30,73	29,24	32,98	31,91	32,36	30,75	28,28	34,98	23,81
26,92	23,51	37,72	33,29	33,87	28,47	35,75	28,16	22,12	34,55
33,41	26,55	27,87	35,60	32,82	33,97	29,79	29,62	31,54	17,39
34,40	32,66	34,88	31,87	26,62	24,35	29,63	29,89	21,17	32,97
26,80	28,88	27,39	30,88	25,20	24,54	28,44	23,73	27,07	20,50
35,32	24,00	33,67	27,30	33,36	25,78	31,49	26,04	28,89	26,41
27,09	25,58	31,19	28,44	28,99	30,20	44,75	34,71	30,94	28,97
31,80	35,66	27,68	35,55	29,07	31,14	29,07	27,08	29,83	36,88
31,51	24,64	33,17	30,10	31,10	43,18	33,52	38,03	29,25	33,19
26,77	23,33	34,26	27,17	33,61	25,92	28,51	35,24	23,17	32,67

Вариант 19

39,40	37,74	23,82	26,07	22,74	23,29	31,48	33,03	40,97	21,97
30,76	36,52	34,25	19,07	36,36	23,26	25,56	37,71	28,31	28,49
41,31	21,28	27,72	29,59	32,26	33,44	31,67	34,28	20,20	32,00
28,17	24,15	32,07	31,09	39,01	34,98	34,75	30,45	29,13	29,11
27,28	28,37	29,47	36,21	26,02	29,70	38,41	36,88	23,53	22,80
36,09	36,19	31,07	25,73	32,59	28,31	24,28	29,37	25,96	32,53
31,39	28,91	26,09	25,30	22,11	37,58	40,43	29,91	26,31	26,13
27,90	29,41	28,68	24,87	27,87	27,90	27,34	26,54	35,18	32,34
31,21	32,82	37,92	35,60	19,48	25,83	25,42	34,22	31,78	36,87
35,20	31,58	36,92	30,81	29,06	28,83	21,89	31,56	28,97	23,11

Вариант 20

25,87	29,92	27,29	40,34	31,24	27,19	25,34	37,37	37,66	32,39
37,17	36,46	27,31	35,15	26,72	32,29	26,78	25,10	34,76	37,94
28,52	27,13	32,51	33,99	32,06	33,43	23,75	34,40	33,17	33,49
35,95	28,55	47,24	35,61	35,87	24,29	24,92	31,55	32,54	28,28
22,88	37,05	18,71	31,71	19,03	23,19	29,03	30,29	15,17	39,18
17,49	30,80	25,18	32,61	30,98	27,51	32,32	33,28	25,54	25,55
18,33	37,32	39,83	19,55	39,36	25,54	31,37	28,63	30,23	25,31
16,13	27,64	27,24	29,51	29,65	38,03	24,72	26,88	34,69	34,38
23,25	27,76	22,55	27,59	24,95	31,98	24,68	41,12	33,67	30,13
36,84	21,92	29,09	37,92	35,52	18,09	26,58	28,92	34,15	36,68

Вариант 21

26,48	26,71	22,48	32,41	20,52	29,33	38,34	25,82	33,75	25,25
21,40	25,13	33,70	28,34	22,71	24,56	23,61	18,13	24,83	38,43
38,99	30,30	30,75	29,89	25,00	30,89	28,93	31,17	28,51	30,76
27,29	33,84	26,75	33,66	38,54	32,69	23,28	26,08	30,43	30,00
34,50	40,94	29,97	32,41	31,60	34,46	36,65	31,15	25,67	32,13
31,77	24,65	33,81	32,84	32,62	29,84	26,02	31,03	32,32	35,87
25,88	40,64	32,85	25,46	29,82	25,83	28,31	32,79	35,43	30,39
36,05	38,48	30,29	34,10	25,93	35,07	36,67	31,29	26,43	33,10
25,33	32,52	33,41	33,91	29,12	37,72	26,90	26,52	31,07	33,97
37,78	20,13	28,33	26,53	35,51	28,66	35,71	33,60	36,78	26,29

Вариант 22

29,17	31,75	21,16	26,34	28,70	23,36	29,14	27,61	28,86	31,52
29,89	26,54	22,85	30,11	35,00	27,05	31,57	35,85	37,48	28,72
29,84	36,47	34,78	25,40	38,53	20,94	34,42	27,84	28,51	29,41
25,24	26,99	24,79	26,68	36,84	23,84	34,29	36,65	21,94	28,19
28,36	22,54	31,69	26,38	31,05	33,45	35,99	32,45	37,55	33,29
26,72	32,82	30,42	38,27	40,00	33,23	35,62	26,81	33,82	33,41
35,41	37,97	31,44	17,84	34,94	31,54	29,01	31,90	31,36	33,45
28,28	28,09	32,61	24,31	28,11	25,73	32,86	33,80	34,57	21,13
28,18	37,20	33,17	29,47	33,41	30,51	28,78	35,41	24,42	23,07
29,41	26,05	22,76	33,91	32,53	25,85	33,32	34,61	25,23	43,93

Вариант 23

28,87	31,03	28,68	34,07	34,54	35,07	30,45	29,27	29,85	44,83
30,02	28,81	30,43	27,67	25,15	33,20	32,70	33,12	28,51	22,54
26,59	34,28	32,06	30,22	31,55	31,21	25,76	23,02	21,62	29,51
24,66	33,08	30,36	29,34	28,90	30,12	32,65	32,94	34,43	29,56
33,80	27,68	30,21	34,83	27,65	23,04	26,73	29,31	20,08	34,74
33,44	27,73	38,07	34,63	35,13	36,03	27,48	31,42	35,11	25,54
32,10	33,35	39,39	35,23	28,44	25,06	25,56	24,82	21,52	36,88
31,21	30,54	29,34	25,96	29,96	25,72	35,02	27,10	31,84	32,50
27,94	35,49	36,20	24,08	24,33	33,61	32,18	27,75	35,17	30,94
27,02	31,76	26,25	25,26	33,34	26,29	32,55	28,79	26,87	36,28

Вариант 24

31,09	30,63	18,79	33,16	24,12	37,58	24,63	20,19	19,78	30,60
26,60	23,26	40,19	30,40	22,57	25,31	23,69	28,03	29,41	33,70
34,90	32,54	25,55	29,64	34,88	32,04	29,97	22,78	22,21	24,51
27,55	23,58	30,00	32,01	25,63	26,13	24,02	36,25	35,06	27,97
28,33	29,29	24,55	36,02	16,50	33,13	30,06	30,45	29,99	31,06
33,95	26,50	34,88	25,66	30,33	25,86	23,74	37,33	29,51	62,23
48,82	33,83	23,77	17,38	37,11	21,63	38,73	27,49	27,44	32,91
36,70	24,18	37,90	27,19	23,48	30,33	30,68	32,44	41,97	31,54
31,70	33,56	39,03	37,58	30,20	40,34	30,87	33,14	28,96	23,99
30,48	24,35	37,97	28,44	30,07	25,86	18,31	42,75	26,03	25,83

Вариант 25

36,51	26,11	31,82	33,69	26,30	28,30	24,98	33,27	28,08	32,13
31,04	29,58	32,20	29,12	30,41	22,83	32,93	27,14	27,79	33,47
23,92	28,92	22,75	31,41	22,82	26,95	31,05	38,83	32,38	25,32
10,76	31,64	29,92	32,25	26,28	36,82	31,26	36,97	35,59	27,56
28,65	37,13	35,20	30,88	31,70	23,25	33,63	23,02	26,99	27,42
30,22	24,93	31,73	29,13	35,66	29,90	32,45	32,69	27,48	25,56
32,36	32,47	32,66	30,30	28,92	28,12	23,67	25,44	28,33	29,74
36,46	31,79	34,36	31,69	31,32	29,76	41,83	29,44	38,42	36,59
26,52	35,29	26,30	36,26	32,23	39,01	29,87	19,27	30,17	38,31
33,96	26,38	31,91	25,05	35,04	32,83	24,77	27,32	19,52	25,68

Практическая работа №2
Построение модели парной регрессии. Оценка параметров модели по методу наименьших квадратов

При эмпирическом (экспериментальном) изучении функциональной зависимости одной величины *y* от другой величины *x* производят ряд измерений величины *y* при различных значениях *x*. Результаты могут быть представлены в виде таблицы 2.1.

Таблица 2.1 – Исходные данные

x	x_1	x_2	…	x_k	…	x_n
y	y_1	y_2	…	y_k	…	y_n

Метод, основанный на требовании минимизации суммы квадратов отклонений, называется методом наименьших квадратов (1МНК).

С его помощью изображают статистическую функциональную зависимость в виде аналитической зависимости и отыскивают такие оценки параметров уравнения регрессии, которые сводят к минимуму выбранную меру разброса. В результате происходит выравнивание эмпирических значений в одну линию регрессии. При этом для однозначного нахождения в качестве меры разброса используют один из показателей рассеяния случайностей величины – дисперсию.

МНК применяется для решения задач, связанных с обработкой результатов испытаний. Этот метод не решает вопроса о выборе вида аналитической функции, а дает возможность в эмпирически подобранной функции определить наиболее вероятные значения для параметров аппроксимирующей функции – в этом и заключается основная задача метода наименьших квадратов.

Пусть после экономического анализа с учетом характера скопления точек $M_i(x_i, y_i)$, $i = \overline{1, n}$, взятых из таблицы 2.1, на плоскости в системе координат *XOY* получена диаграмма рассеяния, по которой и подбирается эмпирическая (аппроксимическая) функция.

Предположим, что диаграмма рассеяния такова, что зависимость между переменными x и y «наилучшим» образом может быть представлена в виде прямой линии $\hat{y} = ax + b$. Это означает, что отклонения фактических значений функции от «подобранной» прямой $Z_i = y_i - \bar{y}_i$ должны быть минимальными, т. е. прямая подбирается так, чтобы сумма квадратов отклонений была наименьшей.

$$z_1^2 + z_2^2 + \ldots + z_i^2 + \ldots z_n^2 \to \min . \qquad (2.1)$$

Пусть

$$\hat{y} = ax + b \qquad (2.2)$$

есть уравнение «подобранной» прямой. Тогда, согласно (2.1), должно выполняться равенство

$$z = \sum_{i=1}^{n} (y_i - ax_i - b)^2 = \min. \qquad (2.3)$$

Требуется определить параметры a и b так, чтобы Z достигла минимума. Известно, что необходимое условие существования минимума состоит в том, чтобы

$$\frac{\partial z}{\partial a} = 0, \frac{\partial z}{\partial b} = 0, \frac{\partial z}{\partial c} = 0. \qquad (2.4)$$

После дифференцирования (2.4) и упрощений получим систему уравнений

$$\begin{cases} a\sum\limits_{i=1}^{n} x_i^2 + b\sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} x_i y_i, \\ a\sum\limits_{i=1}^{n} x_i + bn = \sum\limits_{i=1}^{n} y_i. \end{cases} \qquad (2.5)$$

Для решения системы уравнений (2.5) относительно a и b составляют расчетную таблицу, которая может иметь следующий вид (таблица 2.2).

Таблица 2.2 – Предварительные расчет для решения системы уравнений

i	x_i	y_i	x_i^2	$x_i y_i$
1 2 . . n				
Σ	$\sum\limits_{i=1}^{n} x_i$	$\sum\limits_{i=1}^{n} y_i$	$\sum\limits_{i=1}^{n} x_i^2$	$\sum\limits_{i=1}^{n} x_i y_i$

Полученные в последней строке таблицы суммы подставляем в систему (2.5) и решаем ее любым известным способом.

Система (2.5) называется системой нормальных уравнений в случае выбора эмпирической функции в виде линейной зависимости. Определив параметры a и b, подставляем их в уравнение

$$\bar{y} = ax + b.$$

Если диаграмма рассеяния такова, что эмпирическую зависимость целесообразно выбрать в виде квадратичной функции

$$\hat{y} = ax^2 + bx + c,$$

тогда, согласно МНК, будем иметь

$$z = \sum_{i=1}^{n}(y_i - ax_i^2 - bx_i - c)^2 = \min. \qquad (2.6)$$

В этой функции искомыми величинами являются параметры a, b и c, поэтому согласно необходимым условиям экстремума функции нужно, чтобы

$$\frac{\partial z}{\partial a} = 0, \frac{\partial z}{\partial b} = 0, \frac{\partial z}{\partial c} = 0.$$

Дифференцируя, после упрощений получим

$$\begin{cases} a\sum\limits_{i=1}^{n} x_i^4 + b\sum\limits_{i=1}^{n} x_i^3 + c\sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} x_i^2 y_i, \\ a\sum\limits_{i=1}^{n} x_i^3 + b\sum\limits_{i=1}^{n} x_i^2 + c\sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} x_i y_i, \\ a\sum\limits_{i=1}^{n} x_i^2 + b\sum\limits_{i=1}^{n} x_i + cn = \sum\limits_{i=1}^{n} y_i. \end{cases} \qquad (2.7)$$

Это система нормальных уравнений в случае выбора квадратичной функции в качестве эмпирической функции.

Составим расчетную таблицу 2.3.

Таблица 2.3 – Расчетная таблица для решения системы (2.7)

i	x_i	y_i	x_i^2	x_i^3	x_i^4	$x_i y_i$	$x_i^2 y_i$
1 2 . . n							
$\sum$	$\sum\limits_{i=1}^{n} x_i$	$\sum\limits_{i=1}^{n} y_i$	$\sum\limits_{i=1}^{n} x_i^2$	$\sum\limits_{i=1}^{n} x_i^3$	$\sum\limits_{i=1}^{n} x_i^4$	$\sum\limits_{i=1}^{n} x_i y_i$	$\sum\limits_{i=1}^{n} x_i^2 y_i$

Полученные в последней строке таблицы суммы подставляем в систему (2.7) и решаем ее любым известным способом. Определив параметры a, b и c, подставляем их в уравнение

$$\hat{y} = ax^2 + bx + c\,.$$

Возможен случай, когда эмпирическая квадратичная функция задана в виде

$$\hat{y} = a_0 + a_1 x + a_2 x^2,$$

тогда систему нормальных уравнений можно записать в таком виде:

$$\begin{cases} a_0 n + a_1 \sum\limits_{i=1}^{n} x_i + a_2 \sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} y_i, \\ a_0 \sum\limits_{i=1}^{n} x_i + a_1 \sum\limits_{i=1}^{n} x_i^2 + a_2 \sum\limits_{i=1}^{n} x_i^3 = \sum\limits_{i=1}^{n} x_i y_i, \\ a_0 \sum\limits_{i=1}^{n} x_i^2 + a_1 \sum\limits_{i=1}^{n} x_i^3 + a_2 \sum\limits_{i=1}^{n} x_i^4 = \sum\limits_{i=1}^{n} x_i^2 y_i. \end{cases}$$

Возможен случай, когда диаграмма рассеяния такова, что эмпирическую зависимость целесообразно выбрать в виде гиперболической зависимости

$$\hat{y} = a + \frac{b}{x}. \tag{2.8}$$

Согласно идее МНК нужно, чтобы

$$z = \sum_{i=1}^{n} (y_i - a - \frac{b}{x_i})^2 = \min. \tag{2.9}$$

Нужно подобрать параметры a и b так, чтобы выполнялось условие (2.9), а для этого нужно, чтобы

$$\frac{\partial z}{\partial a} = 0 \quad \text{и} \quad \frac{\partial z}{\partial b} = 0 \tag{2.10}$$

Продифференцировав (2.9) и упростив, получим

$$\begin{cases} an + b\sum\limits_{i=1}^{n} \frac{1}{x_i} = \sum\limits_{i=1}^{n} y_i, \\ a\sum\limits_{i=1}^{n} \frac{1}{x_i} + b\sum\limits_{i=1}^{n} \frac{1}{x_i^2} = \sum\limits_{i=1}^{n} \frac{y_i}{x_i}. \end{cases} \tag{2.11}$$

Это система нормальных уравнений в случае выбора гиперболической функции в качестве эмпирической функции.

Составим расчетную таблицу 2.4.

Таблица 2.4 – Расчетная таблица для решения системы (2.11)

i	x_i	y_i	x_i^2	$\frac{1}{x_i}$	$\frac{1}{x^2{}_i}$	$\frac{y_i}{x_i}$
1 2 . n						
$\sum$	$\sum_{i=1}^{n} x_i$	$\sum_{i=1}^{n} y_i$	$\sum_{i=1}^{n} x_i^2$	$\sum_{i=1}^{n} \frac{1}{x_i}$	$\sum_{i=1}^{n} \frac{1}{x^2{}_i}$	$\sum_{i=1}^{n} \frac{y_i}{x_i}$

Полученные в последней строке таблицы суммы подставляем в систему (2.11) и решаем ее любым известным способом. Определив параметры a и b, подставляем их в уравнение (2.8), найдем эмпирическую функцию.

Может встретиться случай, когда опытные точки $(x_1, y_1),(x_2, y_2),...,(x_n, y_n)$ образуют некоторую линию, похожую на график показательной функции, и эта линия наилучшим образом отражает зависимость между x и y, тогда искомое аппроксимирующее уравнение записывают в виде

$$\hat{y} = a \cdot b^x. \qquad (2.12)$$

Программируем обе части (2.12), получим

$$\log \hat{y} = \log a + x \log b. \qquad (2.13)$$

Введем обозначения: $\log a = A$, $\log b = B$.

Тогда получим уравнение

$$\log \hat{y} = A + Bx. \qquad (2.14)$$

Отсюда следует, что функция $\bar{y}$, представленная на графике в системе координат XOY, где ось ординат разделена по логарифмической шкале, а ось абсцисс – по нормальной шкале, дает прямую с угловым коэффициентом B.

Искомыми параметрами в (2.14) являются A и B. По МНК нужно, чтобы

$$z = \sum_{i=1}^{n} (\log y_i - A - Bx)^2 = \min. \qquad (2.15)$$

Требуется подобрать параметры A и B так, чтобы выполнялось условие (2.15), а для этого нужно, чтобы

$$\frac{\partial z}{\partial a} = 0 \quad \text{и} \quad \frac{\partial z}{\partial b} = 0. \qquad (2.16)$$

Продифференцировав (2.15) и упростив полученное выражение, получим

$$
\begin{cases}
An + B\sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} \log y_i, \\
A\sum\limits_{i=1}^{n} x_i + B\sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} x_i \log y_i.
\end{cases}
\tag{2.17}
$$

Эта система нормальных уравнений в случае выбора функции (2.12) в качестве эмпирической функции.

Составим расчетную таблицу 2.5.

Таблица 2.5 – Расчетная таблица для решения системы (2.17)

i	x_i	y_i	x_i^2	$\log y_i$	$x_i \log y_i$
1 2 . . n					
$\sum$	$\sum\limits_{i=1}^{n} x_i$	$\sum\limits_{i=1}^{n} y_i$	$\sum\limits_{i=1}^{n} x_i^2$	$\sum\limits_{i=1}^{n} \log y_i$	$\sum\limits_{i=1}^{n} x_i \log y_i$

Полученные в последней строке таблицы суммы подставляем в систему (2.17) и решаем ее любым известным способом. Определив параметры A и B, подставляем их в уравнение (2.14). Затем по параметрам A и B пересчитываем интересующие нас параметры a и b.

Следует иметь в виду, что параметры A и B обращают в минимум сумму квадратов отклонений значений преобразованных величин, а не сумму квадратов отклонений измеренных величин y от соответствующих расчетных и могут служить только в качестве первого приближения к наилучшим оценкам отыскиваемых параметров. Поэтому они должны быть уточнены.

Этим способом можно исследовать и такие функции: $\hat{y} = ax^b$; $\bar{y} = ae^{bx}$; в функции $y = \dfrac{1}{ax+b}$ применяют преобразование $\hat{y} = \dfrac{1}{y} = ax + b$ (или $x = \dfrac{1}{x}$, $y = \dfrac{1}{y}$); в функции $\hat{y} = \dfrac{1}{a + be^{-x}}$ применяют преобразование $y = \dfrac{1}{y}$, $x = e^{-x}$.

Эти преобразования необходимы, так как искомые параметры входят в эмпирическую функцию нелинейно.

2.1 Порядок выполнения работы. Пример

Методом наименьших квадратов найти значения параметров эмпирической функции, если опытные данные о значениях X и Y представлены в таблице 2.6.

Таблица 2.6 – Исходные данные задачи

X	1	2	3	4	5	6
Y	15	10	2	2	–4	–10

Решение

По выборке наблюдений построим в системе координат XOY диаграмму рассеяния (рисунок 2.1), т. е. на плоскости в XOY нанесем точки:

$$(1,15) \quad (2,10) \quad (3,2) \quad (4,2) \quad (5,-4) \quad (6,-10)$$

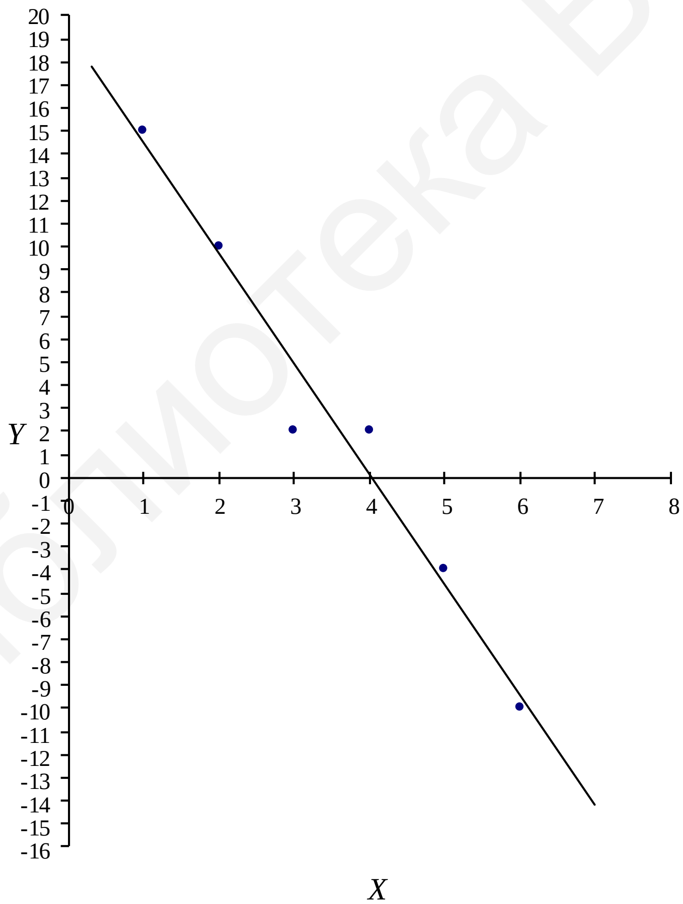


Рисунок 2.1 – Диаграмма рассеяния

Анализ опытных данных показывает, что в качестве эмпирической (подобранной) функции можно использовать линейную функцию

$$\hat{y} = a \cdot x + b. \qquad (2.18)$$

В выражении (2.18) необходимо найти параметры a и b, для чего применяем МНК. Тогда для нахождения параметров будем иметь систему нормальных уравнений

$$\begin{cases} a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i, \\ a\sum_{i=1}^{n} x_i + b \cdot n = \sum_{i=1}^{n} y_i. \end{cases}$$

Для удобства вычислений составим следующую расчетную таблицу 2.7 (из условия задачи известно, что $n = 6$).

Таблица 2.7 – Расчетная таблица для решения задачи

i	x_i	y_i	x_i^2	$x_i \cdot y_i$
1	1	15	1	15
2	2	10	4	20
3	3	2	9	6
4	4	2	16	8
5	5	–4	25	–20
6	6	–10	36	–60
$\sum$	$\sum_{i=1}^{6} x_i = 21$	$\sum_{i=1}^{6} y_i = 15$	$\sum_{i=1}^{6} x_i^2 = 91$	$\sum_{i=1}^{6} x_i \cdot y_i = -31$

Подставим данные последней строки таблицы в систему уравнений (2.5):

$$\begin{cases} 91 \cdot a + 21 \cdot b = -31, \\ 21 \cdot a + 6 \cdot b = 15. \end{cases}$$

Решая эту систему любым известным способом, получим

$$a = -4{,}76; \; b = 19{,}2.$$

Подставляя эти значения параметров в (2.18), получим эмпирическую функцию

$$\hat{y} = -4{,}76 \cdot x + 19{,}2.$$

Если эмпирическая функция будет записана в виде $\hat{y} = a + b \cdot x$, то система нормальных уравнений имеет вид

$$\begin{cases} a \cdot n + b\sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} y_i, \\ a \cdot \sum\limits_{i=1}^{n} x_i + b \cdot \sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} x_i y_i. \end{cases} \tag{2.19}$$

Рассмотрим пример применения МНК в матричной форме.

2.2 Варианты заданий

По заданным статистическим данным подобрать эмпирическую функцию, если она не задана, и выполнить следующие действия:

- построить диаграмму рассеяния;
- записать эмпирическую функцию;
- записать систему нормальных уравнений;
- составить расчетную таблицу;
- решить полученную систему и записать эмпирическую функцию с найденными параметрами.

1 Считая, что зависимость между переменными x и y имеет вид $y = ax + b$, найти оценки параметров по следующим выборкам.

Варианты	**Задания**										
1	x	54	63	74	90	112	140	190	189	210	230
	y	8	10	11	13	15	17	19	21	23	25
2	x	100	120	110	115	125	130	125	140	140	150
	y	12	13	18	19	20	20	25	30	31	35
3	x	1	3	4	2	5	7	8	9	10	11
	y	80	90	120	100	110	150	160	130	126,66	116,6
4	x	5	4	6	7	3	4	6	7	4	3
	y	6,3	6	7,5	8,5	3,5	6,2	7,5	8,7	6	3,7
5	x	152	116	100	108	129	141	147	156	163	171,333
	y	47	34	31	32	38	42	45	47	49	52
6	x	90	110	120	130	180	200	280	360	440	520
	y	25	28	31	32	36	42	55	68	81	94
7	x	2	4	3	5	2	2	5	4	6	4
	y	13	15	12	16	15	11	14	11	13	12
8	x	6	6,1	6,8	7,2	7,4	7,9	8,2	8,5	8,6	9,1
	y	2	3	6	4	3	3	4	5	6	8
9	x	6	8	9	9	10	11	11	13	14	15
	y	4	4	5	7	5	6	8	7	9	10
10	x	8	9	10	11	12	13	14	16	17	19
	y	9	8,5	9,2	9,6	9,4	10,5	11,2	10,8	11	11,5
11	x	2	2	3	4	5	6	7	8	9	11
	y	2,5	3,1	3	3,5	4,2	5,1	5,5	6	6,2	6,4

Варианты	Задания										
12	*x*	66	70	75	80	82	85	90	92	95	98
	y	60	68	65	78	74	70	78	85	88	90
13	*x*	13	14	15	17	18	18	18	19	22	25
	y	7	9	10	12	11	14	15	15	16	18
14	*x*	0	4	10	15	21	29	36	51	68	72
	y	66,7	71	76,3	80,6	85,7	92,9	99,4	113,6	125,1	130
15	*x*	0,3	0,91	1,5	2	2,2	2,62	3	3,3	3,5	3,8
	y	0,2	0,43	0,35	0,52	0,81	0,68	1,15	0,85	0,82	0,9
16	*x*	37	47	49	51	61	75	80	92	102	117
	y	53	42	30	24	22	22	26	31	35	38
17	*x*	1	2	3	4	5	6	7	8	9	10
	y	7,6	7,2	6,2	8,3	8,2	7,6	7,9	7,5	8,5	8,7
18	*x*	1	5	6	7	8	9	10	11	12	13
	y	100	156	170	184	194	205	220	229	238	247
19	*x*	1	2	3	4	5	6	7	8	9	10
	y	100	113	121	148	183	194	219	260	277	304
20	*x*	2	5	8	10	14	15	4	12	3	7
	y	14,39	9,45	7,05	5,32	16,94	1,97	8,75	3,41	13,37	8,22
21	*x*	2,7	4,6	6,3	7,8	9,2	10,6	12	13,4	14,8	16,2
	y	17	16,2	13,3	13	9,7	9,9	6,2	5,8	5,4	5
22	*x*	7,9	11,6	12,01	14,9	16,3	18,6	20,30	21,9	23,6	25,23
	y	13	22,8	24,8	28,6	31,6	38,7	40	44,9	43	45,63
23	*x*	1	2	3	4	5	6	7	8	9	10
	y	0,21	0,32	0,58	1,02	1,76	2,68	3,75	5,07	6,62	8,32
24	*x*	2	4	6	8	10	12	14	16	18	20
	y	4,5	7	8	7,5	9	10,5	12	13,5	15	16,5

2 Считая, что зависимость между переменными x и y имеет вид $y = ax^2 + bx + c$, найти оценки параметров по следующим выборкам.

Вариант 1

x	2,0	3,5	4,0	4,5	5,5	6,0
y	1,9	1,7	1,8	1,6	1,5	1,4

Вариант 2

x	5,0	6,0	6,5	7,0	8,0
y	25	2,8	31	35	40

Вариант 3

x	2	3	4	4	5	6	6	6	7	8
y	8	10	7	6	5	5	4	3	4	5

Вариант 4

x	40	55	64	75	82	94	104	110	115	120
y	2,8	4,3	4,6	4,9	5,6	6,4	7,7	7,9	10,2	9,8

Вариант 5

x	10	20	30	40	50	60	70	80	90
y	4,2	12,6	14,8	16,8	21,0	22,2	22,8	21,8	19,4

Вариант 6

x	7	12	17	22	27	32	37
y	83,7	72,9	63,2	54,7	47,5	41,4	36,3

Вариант 7

x	12,0	13,1	14,0	16,1	17,4	18,0
y	54	59	67	76	85	97
x	20,0	21,4	21,9	24,1	25,0	26,8
y	107	118	127	139	153	160

Вариант 8

x	–2	–1	0	1	2	3
y	–3	–4	–3	0	5	12

Вариант 9

x	1	2	3	4	5	6	7
y	0,5	0,5	1	2	3	5	8

В последующих примерах взять эмпирическую функцию в виде

$$y = a_0 + a_1 x + a_2 x^2 .$$

Вариант 10

x	0	2	4	6	8	10
y	5	–1	–0,5	1,5	4,5	8,5

Вариант 11

x	0,07	0,31	0,61	0,99	1,29	1,78	2,09
y	1,34	1,08	0,94	1,06	1,25	2,01	2,6

Вариант 12

x	26	30	34	38	42	46	50
y	3,94	4,60	5,67	6,93	8,25	7,73	10,55

Вариант 13

x	–2	–1	0	1	2
y	4,8	0,4	–3,4	0,8	3,2

Вариант 14

x	–3	–2	–1	0	1	2	3
y	–10	0	4	5	4	2	–2

Вариант 15

x	–3	–2	–1	0	1	2	3
y	–5	–2	–1,5	–1	0	3	14

Вариант 16

x	–3	–2	–1	0	1	2	3
y	6	0	–1	–1	1	5	12

Вариант 17

x	–2	–1	0	1	2
y	3	0	3	6	9

Вариант 18

x	–3	–2	–1	0	2	3
y	–6	–4	–2	–1	1	0

Вариант 19

x	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0	5,5	6,0
y	0,4	0,3	1,0	1,7	2,1	3,4	4,1	5,8	7,7	9,4	11,4	13,6

Вариант 20

x	0,4	0,8	1,2	1,6	2	2,4	2,8
y	0,43	0,94	1,91	3,01	4	4,56	6,45
x	3,2	3,6	4,0	4,4	4,8	5,2	–
y	8,59	11,15	13,88	16,93	20,47	24,15	–

Вариант 21

x	0	0,05	0,10	0,15	0,20	0,25	0,30
y	25	26	4	7	6	13	30
x	0,35	0,40	0,45	0,50	0,55	0,60	–
y	26	32	40	32	21	11	–

Вариант 22

x	1,0	1,5	2,0	2,5	3,0	3,5
y	0,22	0,23	0,31	0,43	0,56	0,82
x	4,0	4,5	5,0	5,5	6,0	6,5
y	1,06	1,25	1,72	2,28	2,67	3,26

Вариант 23

x	5	10	15	20	25
y	50	59,3	59,8	64,9	70,2

Вариант 24

x	87,5	84,0	77,8	63,7	46,7	36,9
y	292	283	270	235	197	181

3 Считая, что зависимость x и y имеет вид $y = a + \frac{b}{x}$, найти оценки параметров по следующим выборкам.

Вариант 1

x	1	2	3	4	5
y	16,50	13,75	13,31	12,50	13,52
x	6	7	8	9	10
y	12,75	12,30	12,83	12,28	12,34

Вариант 2

x	1	2	3	4	5
y	1,25	1,15	1,00	0,80	0,65
x	10	20	30	50	100
y	0,41	0,36	0,20	0,15	0,1

Вариант 3

x	54	63	74	90	112	140	190
y	0,50	0,70	0,80	1,00	1,40	2,20	2,50

Вариант 4

x	40	55	64	75	82	94	104	110	115	120
y	2,8	4,3	4,6	4,9	5,6	6,4	7,7	7,9	10,2	9,8

Вариант 5

x	54	63	74	90	112	140	190
y	0,15	0,20	0,35	0,55	0,95	1,40	1,60

Вариант 6

x	1,0	0,5	0,07	0,3	0,25	0,34	0,13	0,08	0,22	0,58
y	1,6	1,0	8,5	5,0	4,4	2,0	6,0	7,5	3,8	1,4

Вариант 7

x	75	90	120	150	180	220	300	450	600	800
y	10	9,2	8,1	7,8	7,9	7,0	6,1	5,8	5,3	5,0

Вариант 8

x	5,67	4,45	3,84	3,74	3,73	2,18
y	6,8	8,5	10,5	10,2	6,8	11,8

Вариант 9

x	2	4	6	12
y	8	5,25	3,50	3,25

Вариант 10

x	1	2	3	4	5
y	16,50	13,75	13,31	12,50	13,52
x	6	7	8	9	10
y	12,75	12,30	12,83	12,28	12,34

Вариант 11

x	2,6	5,2	7,8	15,6
y	8	5,25	3,50	3,25

Вариант 12

x	5,67	4,45	3,84	3,74	3,73	2,18
y	6,8	8,5	10,5	10,2	6,8	11,8

Вариант 13

x	3	6	9	18
y	8	5,25	3,50	3,25

Вариант 14

x	1	2	3	5	10	20	30	50	100	200
y	10,15	5,52	4,08	2,85	2,11	1,62	1,41	1,30	1,21	1,15

Вариант 15

x	8,8	11	13,2	14,85	15,4
y	80	72	65	70	68

Вариант 16

x	8,1	16,1	21,8	43,9	65,8	87,6	96,5
y	0,330	0,271	0,242	0,183	0,158	0,142	0,138

Вариант 17

x	8	10	12	13,5	14
y	80	72	65	70	68

Вариант 18

x	0,125	0,25	0,5	1	2	4	8
y	170	90	50	30	20	15	12,5

Вариант 19

x	5	7	9	11	13	15	17	19
y	3,6	4,0	4,4	4,8	5,2	5,6	6,0	6,4

Вариант 20

x	5	7	9	11	13	15	17	19
y	6,16	5,31	4,83	4,53	4,32	4,17	4,05	3,96

Вариант 21

x	37,8	44,1	51,8	63	78,4	98	133
y	0,15	0,20	0,35	0,55	0,95	1,40	1,60

Вариант 22

x	37,5	45	60	75	90	110	150	225	300	400
y	10	9,2	8,1	7,8	7,9	7,0	6,1	5,8	5,3	5,0

Вариант 23

x	3,5	7	10,5	14	17,5
y	59,3	59,8	60,1	64,9	70,2

Вариант 24

x	37,8	44,1	51,8	63,0	78,4	98	133
y	0,50	0,70	0,80	1,0	1,40	2,20	2,50

Практическая работа №3
Построение линейной многофакторной эконометрической модели

3.1 Оценка параметров линейной многофакторной эконометрической модели

Многофакторная линейная эконометрическая модель (множественная линейная регрессионная модель) устанавливает линейную зависимость между одним показателем и несколькими факторами.

В общем виде многофакторная линейная эконометрическая модель записывается следующим образом:

$$y = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n + e, \tag{3.1}$$

где y – вектор наблюдений за результирующим показателем;

$x_1, x_2, \ldots, x_n$ – вектор наблюдений за факторами;

$a_1, a_2, \ldots, a_n$ – вектор неизвестных параметров модели;

e – вектор случайных величин.

Многофакторную линейную эконометрическую модель целесообразно изучать с использованием теории матриц. Уравнение (3.1) в матричной форме запишется следующим образом:

$$\begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1m} \\ 1 & x_{21} & x_{22} & \ldots & x_{2m} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_{n1} & x_{n2} & \ldots & x_{nm} \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \ldots \\ a_m \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \ldots \\ e_n \end{pmatrix} \tag{3.2}$$

или $Y = X \cdot A + e$.

Оценкой модели будет модель

$$\hat{Y} = X \cdot \hat{A}, \tag{3.3}$$

где Y – вектор-столбец наблюдаемых значений показателя;

$\hat{Y}$ – вектор-столбец оцененных значений показателя y;

X – матрица наблюдаемых значений факторов;

X_0 – фиктивный фактор, все значения которого равны единице;

A – вектор-столбец неизвестных параметров;

$\hat{A}$ – вектор-столбец оценок параметров;

e – вектор-столбец остатков (отклонений).

Оценка параметров модели осуществляется методом 1МНК при условии, что все допущения классической регрессионной модели выполняются. Оператор оценки параметров модели 1МНК в матричном виде

$$\hat{A} = (X^T \cdot X)^{-1} \cdot X^T \cdot Y \tag{3.4}$$

где $X^T \cdot X$ – матрица моментов;

$(X^T \cdot X)^{-1}$ – матрица ошибок.

3.2 Ковариационная и кореляционная матрицы

Взаимосвязь между отдельными признаками описывается при помощи ковариационной или корреляционной матрицы. Матрица, составленная из ковариаций и дисперсий, называется ковариационной матрицей и имеет вид

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1m} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2m} \\ \dots & \dots & \dots & \dots \\ \sigma_{m1} & \sigma_{m2} & \dots & \sigma_m^2 \end{pmatrix}. \qquad (3.5)$$

Свойства ковариационной матрицы:

- ковариационная матрица симметрична;
- на главной диагонали размещены дисперсии признаков $X_j - \sigma_j^2 \ \left(j = \overline{1;m}\right)$;
- вне главной диагонали – ковариации признаков X_i и $X_j - \sigma_j^2 \ \left(j = \overline{1;m},\ i = \overline{1;m}\right)$.

Матрица, составленная из коэффициентов парной корреляции, называется корреляционной матрицей и имеет вид

$$r = \begin{pmatrix} 1 & r_{12} & \dots & r_{1m} \\ r_{21} & 1 & \dots & r_{2m} \\ \dots & \dots & \dots & \dots \\ r_{m1} & r_{m2} & \dots & 1 \end{pmatrix}. \qquad (3.6)$$

Свойства корреляционной матрицы:

- корреляционная матрица симметрична;
- на главной диагонали размещены единицы;
- вне главной диагонали – коэффициенты парной корреляции признаков X_i и $X_j - r_{ij} \ \left(j = \overline{1;m},\ i = \overline{1;m}\right)$.

При построении двухфакторной модели определяется корреляционная матрица, которая учитывает взаимосвязь между показателем и фактором в виде

$$r = \begin{pmatrix} 1 & r_{yx_1} & r_{yx_2} \\ r_{yx_1} & 1 & r_{x_1x_2} \\ r_{yx_2} & r_{x_1x_2} & 1 \end{pmatrix}. \tag{3.7}$$

Коэффициент парной корреляции между показателями и каждым фактором, а также между факторами рассчитывается по формулам:

$$r_{x_1y} = \frac{\sigma_{x_1y}}{\sigma_{x_1}\sigma_y}, \quad r_{x_2y} = \frac{\sigma_{x_2y}}{\sigma_{x_2}\sigma_y}, \quad r_{x_1x_2} = \frac{\sigma_{x_1x_2}}{\sigma_{x_1}\sigma_{x_2}}, \tag{3.8}$$

где $\sigma_y = \sqrt{\sigma_y^2}$ – среднее квадратическое отклонение показателя Y;

$\sigma_{x_1} = \sqrt{\sigma_{x_1}^2}$ – среднее квадратическое отклонение фактора X_1;

$\sigma_{x_2} = \sqrt{\sigma_{x_2}^2}$ – среднее квадратическое отклонение фактора X_2;

$\sigma_y^2 = \overline{y^2} - \bar{y}^2$ – дисперсия показателя Y;

$\sigma_{x_1}^2 = \overline{x_1^2} - \bar{x}_1^2$ – дисперсия фактора X_1;

$\sigma_{x_2}^2 = \overline{x_2^2} - \bar{x}_2^2$ – дисперсия фактора X_2;

$\sigma_{x_1y} = \overline{x_1y} - \bar{x}_1 \cdot \bar{y}$ – коэффициент ковариации показателя X_1 и показателя Y;

$\sigma_{x_2y} = \overline{x_2y} - \bar{x}_2 \cdot \bar{y}$ – коэффициент ковариации фактора X_2 и показателя Y;

$\sigma_{x_1x_2} = \overline{x_1x_2} - \bar{x}_1 \cdot \bar{x}_2$ – коэффициент ковариации факторов X_1 и X_2.

3.3 Коэффициент частичной корреляции

Частичной корреляцией между признаками X_i и X_j называется корреляционная зависимость между этими признаками при фиксированных значениях других признаков.

Формула частичного коэффициента между признаками X_i и X_j имеет вид

$$r_{ij(1,2,\ldots,m)} = \frac{-R_{ij}}{\sqrt{R_{ii} \cdot R_{jj}}}, \tag{3.9}$$

где R_{ij}, R_{ii}, R_{jj} – алгебраические дополнения соответствующих элементов матрицы R.

При построении двухфакторной модели коэффициенты частичной корреляции рассчитываются по формулам

$$r_{yx_1(x_2)} = \frac{r_{yx_1} - r_{yx_2} r_{x_1x_2}}{\sqrt{(1 - r^2_{yx_2})(1 - r^2_{x_1x_2})}},$$

$$r_{yx_2(x_1)} = \frac{r_{yx_2} - r_{yx_1} r_{x_1x_2}}{\sqrt{(1 - r^2_{yx_1})(1 - r^2_{x_1x_2})}}, \qquad (3.10)$$

$$r_{x_1x_2(y)} = \frac{r_{x_1x_2} - r_{yx_1} r_{yx_2}}{\sqrt{(1 - r^2_{yx_1})(1 - r^2_{yx_2})}}.$$

При использовании электронных талиц Excel расчеты целесообразно выполнить по формуле коэффициента частичной корреляции через элементы матрицы, обратной корреляционной матрице, которая имеет вид

$$r_{ij(1,2,\ldots,m)} = \frac{-c_{ij}}{\sqrt{c_{ii} \cdot c_{jj}}}, \qquad (3.11)$$

где c_{ij}, c_{ii}, c_{jj} – элементы матрицы $C = (R)^{-1}$, обратной корреляционной матрице R.

3.4 Коэффициенты множественной корреляции и детерминации. ANOVA-дисперсионный анализ

Коэффициент множественной корреляции является мерой степени соответствия оцененной модели фактическим данным и определяется как коэффициент корреляции между y и $\hat{y}$.

Этот коэффициент может использоваться при любой форме связи между признаками.

Квадрат коэффициента множественной корреляции называется коэффициентом множественной детерминации. Коэффициент множественной детерминации характеризует долю дисперсии (вариации) показателя y, что объясняется регрессией, т. е. вариацией факторов, входящих в модели:

$$R^2 = \frac{SSR}{SST},$$

где R^2 – коэффициент множественной детерминации;

$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$ – сумма квадратов общих отклонений;

$SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$ – сумма квадратов отклонений, которые объясняются регрессией.

Учитывая, что $SST = SSR + SSE$, где $SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ – сумма квадратов отклонений, которые не объясняются регрессией, получаем формулу

$$R^2 = 1 - \frac{SSE}{SST} \quad \text{или} \quad R^2 = 1 - \frac{\sum_{i-1}^{n} e_i^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}. \qquad (3.12)$$

В матричной форме коэффициент множественной детерминации имеет вид

$$R^2 = \frac{\hat{A}^T \cdot X^T \cdot Y - n \cdot \bar{y}^2}{Y^T \cdot Y - n \cdot \bar{y}^2}. \qquad (3.13)$$

Коэффициент множественной корреляции удобно рассчитывать как квадратный корень из коэффициента множественной детерминации, т. е. $R = \sqrt{R^2}$.

При использовании электронных таблиц Excel во время проведения дисперсионного анализа целесообразно проанализировать таблицу ANOVA (Analysis of Variance – дисперсионный анализ) пакета Анализ данных (таблица 3.1).

Таблица 3.1 – ANOVA-дисперсионный анализ

Наименование	**Степени свободы**	**Сумма квадратов**	**Средние квадраты**
Обозначение	df	SS	MS
Регрессия	m	SSR	$MSR = SSR / m = \hat{\sigma}_{\hat{y}}^2$
Остаток	$n - (m + 1)$	SSE	$MSE = SSE / (n - m - 1) = \hat{\sigma}_e^2$
Вместе	$n - 1$	SST	$MST = SST / (n - 1) = \hat{\sigma}_y^2$

Примечание – В таблице 3.1 использованы следующие обозначения:

m – количество факторов модели;

n – количество наблюдений;

$\hat{\sigma}_{\hat{y}}^2$ – оценка дисперсии оцененных значений показателя;

$\hat{\sigma}_e^2$ – оценка дисперсии отклонений;

$\hat{\sigma}_y^2$ – оценка дисперсии показателя.

Корень квадратный из оценки дисперсии отклонений называется стандартной ошибкой оценки:

$$\hat{\sigma}_e = \sqrt{\hat{\sigma}_e^2}\,. \qquad (3.14)$$

Для уравнивания моделей с разным количеством факторов необходимо использовать нормированный (скорректированный) коэффициент детерминации:

$$R^2{}_{\text{норм}} = 1 - \frac{MSE}{MST} \text{ или } R^2{}_{\text{норм}} = 1 - \frac{\hat{\sigma}_e^2}{\hat{\sigma}_y^2}\,. \qquad (3.15)$$

Примечание – При использовании электронных таблиц Excel значения коэффициентов множественной корреляции и детерминации, нормированого коэффициента детерминации и стандартной ошибки можно найти в регрессионной статистике вывода итогов пакета Анализ данных.

3.5 Разложение коэффициента множественной детерминации на коэффициенты отдельной детерминации

Для выяснения доли влияния каждого фактора на показатель используются коэффициенты отдельной детерминации.

Коэффициентом отдельной детерминации $d_j^{\,2}$ для фактора X_j называется произведение коэффициента корреляции $r_{x_j y}$ между фактором X_j и показателем Y и стандартизированного параметра регрессии a_j^*.

$$d_j^2 = r_{yx_j} \cdot a_j^*, \qquad (3.16)$$

где $a_j^* = \dfrac{\hat{a}_j \cdot \sigma_{x_j}}{\sigma_y}$ – стандартизированный параметр регрессии.

Сумма коэффициентов отдельной детерминации равна коэффициенту множественной детерминации, т. е.

$$R^2{}_{\text{норм}} = \sum_{j=1}^{m} d_j^2\,. \qquad (3.17)$$

При анализе двухфакторной модели коэффициенты отдельной детерминации рассчитываются по формулам:

$$d_1^2 = r_{yx_1} \cdot a_1^* = r_{yx_1} \frac{\hat{a}_1 \cdot \sigma_{x_1}}{\sigma_y},$$

$$d_2^2 = r_{yx_2} \cdot a_2^* = r_{yx_2} \frac{\hat{a}_2 \cdot \sigma_{x_2}}{\sigma_y}, \qquad (3.18)$$

$$R^2 = d_1^2 + d_2^2.$$

Примечание – Коэффициент парной корреляции $r_{x_j y}$ между фактором X_j и показателем Y в многомерной модели измеряет нечистое влияние этого фактора на показатель, а стандартизированный параметр регрессии a_j^* измеряет условно чистое влияние фактора па показатель. Поэтому отдельные случайные значения $r_{x_j y}$ и a_j^* могут иметь разные знаки, в этом случае появляются неинтерпретированные отрицательные величины коэффициентов отдельной детерминации.

3.6 Ковариационная матрица оценок параметров модели

Оцененная ковариационная матрица оценок параметров модели состоит из оценок ковариаций и дисперсий оценок параметров модели и вычисляется по формуле

$$\hat{\Sigma}_{\hat{A}} = \hat{\sigma}_e^2 \cdot (X^T \cdot X)^{-1} = \begin{pmatrix} \hat{\sigma}_{\hat{a}_0}^2 & \hat{\sigma}_{\hat{a}_0 \hat{a}_1} & \ldots & \hat{\sigma}_{\hat{a}_0 \hat{a}_m} \\ \hat{\sigma}_{\hat{a}_1 \hat{a}_0} & \hat{\sigma}_{\hat{a}_1}^2 & \ldots & \hat{\sigma}_{\hat{a}_1 \hat{a}_m} \\ \ldots & \ldots & \ldots & \ldots \\ \hat{\sigma}_{\hat{a}_m \hat{a}_0} & \hat{\sigma}_{\hat{a}_m \hat{a}_1} & \ldots & \hat{\sigma}_{\hat{a}_m}^2 \end{pmatrix}, \qquad (3.19)$$

где $\hat{\sigma}_e^2 = \dfrac{\sum_{i=1}^{n} e_i^2}{n-m-1}$ – оценка дисперсии отклонений.

Квадратный корень из оценки дисперсии оценок параметров модели называется стандартной ошибкой оценок параметров модели, т. е.

$$\hat{\sigma}_{\hat{a}_j} = \sqrt{\hat{\sigma}_{\hat{a}_j}^2}, \qquad j = \overline{0;m}. \qquad (3.20)$$

При анализе двухфакторной модели стандартные ошибки оценок параметров модели рассчитываются по формулам:

$$\hat{\sigma}_{\hat{a}_0} = \sqrt{\hat{\sigma}_{\hat{a}_0}^2}, \qquad \hat{\sigma}_{\hat{a}_1} = \sqrt{\hat{\sigma}_{\hat{a}_1}^2}, \qquad \hat{\sigma}_{\hat{a}_2} = \sqrt{\hat{\sigma}_{\hat{a}_2}^2}. \qquad (3.21)$$

Если стандартная ошибка превышает 10 % от абсолютного значения параметра модели, то это свидетельствует о смещении оценки этого параметра, т. е.:

$$\text{если } \frac{\hat{\sigma}_{\hat{a}_j}}{|\hat{a}_j|} \cdot 100\,\% > 10\,\%, \quad \text{то } \hat{a}_j - \text{смещенная оценка}, \quad j = \overline{0;m}. \qquad (3.22)$$

Этот факт объясняется неточной спецификацией модели: в модель не включен значимый фактор или неверно выбрана аналитическая зависимость.

Примечание – При использовании электронных таблиц Excel расчетные значения стандартных ошибок оценок параметров модели можно найти в таблице ВЫВОД ИТОГОВ пакета АНАЛИЗ ДАННЫХ – РЕГРЕССИЯ.

3.7 Проверка модели на адекватность по F-критерию Фишера

Наиболее распространенным из статистических критериев, которые однозначно дают ответ на вопрос об адекватности построенной модели, является критерий Фишера. Проверку модели на адекватность по *F*-критерию Фишера целесообразно осуществлять по следующему алгоритму.

Алгоритм тестирования по F-критерию Фишера

Шаг 1. Формулировка нулевой и альтернативной гипотез.

H_0: $\hat{a}_0 = \hat{a}_1 = \ldots = \hat{a}_m = 0$, т. е. ни один фактор модели не влияет на показатель, или все параметры модели незначимы.

H_A: хотя бы одно значение $\hat{a}_j$, отличное от нуля, т. е. $\hat{a}_j \neq 0$, $j = \overline{0;m}$.

Шаг 2. Выбор подходящего уровня значимости.

Уровнем значимости α называется вероятность сделать ошибку 1-го рода, т. е. отклонить правильную гипотезу. Величина $P = 1 - \alpha$ называется уровнем доверия или доверительной вероятностью.

В электронных таблицах Excel по умолчанию принимается $\alpha = 0{,}05\,(P = 0{,}95)$. Это означает, что мы рискуем отклонить правильную гипотезу в 5 % случаев, а в 95 % случаев наши выводы будут верными.

Шаг 3. Расчет расчетного значения F-критерия.

Расчетное значение *F*-критерия, так называемое *F*-отношение, определяется по формуле

$$F_{\text{расч}} = \frac{MSR}{MSE}, \qquad (3.23)$$

где $F_{\text{расч}}$ – расчетное значение *F*-критерия;

MSR – средний квадрат отклонений, обусловленных регрессией;

MSE – средний квадрат отклонений, которые не объясняются регрессией (см. таблицу 3.1). Или по формуле

$$F_{\text{расч}} = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m}, \qquad (3.24)$$

где R^2 – коэффициент множественной детерминации.

Примечание – При использовании электронных таблиц Excel расчетное значение F-критерия можно найти в таблице *Дисперсионный анализ* ВЫВОДА ИТОГОВ пакета АНАЛИЗ ДАННЫХ – РЕГРЕССИЯ.

Шаг 4. Определение по статистическим таблицам F-распределения Фишера критического значения F-критерия.

Критическое значение *F*-критерия находят по статистическим таблицам *F*-распределения Фишера по соответствующим значениям:

- доверительной вероятности Р;
- степеней свободы $k_1 = m$ и $k_2 = n - m - 1$.

Шаг 5. Сравнение расчетного значения F-критерия с критическим и интерпретация результатов тестирования.

Если $F_{\text{расч}} < F_{\text{кр}}$, то нет оснований отклонить нулевую гипотезу о том, что ни один фактор модели не является значимым, т. е. с принятой надежностью можно утверждать, что модель неадекватна.

Если $F_{\text{расч}} > F_{\text{кр}}$, то нулевая гипотеза о незначительности факторов отклоняется, т. е. с принятой надежностью можно утверждать, что модель адекватна.

3.8 Проверка значимости оценок параметров модели по t-критерию Стьюдента

Т-распределение Стьюдента позволяет протестировать гипотезы о значимости каждого параметра модели и построить их интервалы доверия.

Алгоритм тестирования по t-критерию Стьюдента

Шаг 1. Формулировка нулевой и альтернативной гипотез.

H_0: $\hat{a}_0 = 0, \quad j = \overline{0;\ 2}$, – оценка *j*-го параметра является статистически незначимой, и *j*-й фактор никак не влияет на показатель *y*;

H_A: $\hat{a}_0 \neq 0, \quad j = \overline{0;\ 2}$, – оценка *j*-го параметра является статистически значимой, и *j*-й фактор влияет на показатель *y*.

Шаг 2. Выбор подходящего уровня значимости.

Уровень значимости избирается аналогично *F*-критерию.

Шаг 3. Расчет расчетного значения t-критерия.

Расчетные значения *t*-критерия определяются по формулам:

$$\hat{t}_{\hat{a}_j} = \frac{\hat{a}_j}{\hat{\sigma}_{\hat{a}_j}}, \quad j = \overline{0;m}. \tag{3.25}$$

При анализе двухфакторной модели расчетные значения *t*-критерия определяются по формулам:

$$\hat{t}_{\hat{a}_0} = \frac{\hat{a}_0}{\hat{\sigma}_{\hat{a}_0}}, \qquad \hat{t}_{\hat{a}_1} = \frac{\hat{a}_1}{\hat{\sigma}_{\hat{a}_1}}, \qquad \hat{t}_{\hat{a}_2} = \frac{\hat{a}_2}{\hat{\sigma}_{\hat{a}_2}}. \tag{3.26}$$

Примечание – При использовании электронных таблиц Excel расчетные значения *t*-критерия можно найти в таблице ВЫВОД ИТОГОВ пакета АНАЛИЗ ДАННЫХ – РЕГРЕССИЯ.

Шаг 4. Определение по статистическим таблицам t-распределения Стьюдента критического значения t-критерия.

Критическое значение *t*-критерия находят по статистическим таблицам *t*-распределения Стьюдента по соответствующим значениям:

- доверительной вероятности *P*;
- числа степеней свободы $k = n - m - 1$.

Шаг 5. Сравнение расчетного значения t-критерия с критическим и интерпретация результатов тестирования.

Если $|t_{\text{расч}}| < t_{\text{кр}}$, то нет оснований отклонять нулевую гипотезу, т. е. с принятой надежностью можно утверждать, что оценка *j*-го параметра является статистически незначимой, *j*-й фактор не влияет на показатель *y*.

Если $|t_{\text{расч}}| > t_{\text{кр}}$, то нулевая гипотеза отклоняется, т. е. с принятой надежностью можно утверждать, что оценка *j*-го параметра является статистически значимой, *j*-й фактор влияет на показатель *y*.

3.9 Определения интервала доверия для параметров модели

Интервалом доверия называется интервал, который содержит неизвестный параметр с заданным уровнем доверия.

Процедура нахождения интервалов доверия для параметров аналогична процедуре тестирования нуль-гипотезы по *t*-критерию Стьюдента:

- избирается уровень значимости α, соответственно уровень доверия $P = 1 - \alpha$;
- определяется критическое значение *t*-критерия $t_{\alpha/2}$ по статистическим таблицам *t*-распределения Стьюдента;
- для каждого параметра вычисляются нижняя и верхняя границы интервалов доверия по формулам:

$$\hat{a}_j - t_{\alpha/2}\hat{\sigma}_{\hat{a}_j} \le a_j \le \hat{a}_j + t_{\alpha/2}\hat{\sigma}_{\hat{a}_j}, \quad j = \overline{0;m}, \tag{3.27}$$

где $\hat{\sigma}_{\hat{a}_j}$ – стандартные ошибки оценок параметров модели $\hat{a}_j$, $j = \overline{0;m}$.

При анализе двухфакторной модели доверительные интервалы для параметров определяются по формулам:

$$\begin{aligned} &\hat{a}_0 - t_{\alpha/2}\hat{\sigma}_{\hat{a}_0} \le a_0 \le \hat{a}_0 + t_{\alpha/2}\hat{\sigma}_{\hat{a}_0}, \\ &\hat{a}_1 - t_{\alpha/2}\hat{\sigma}_{\hat{a}_1} \le a_1 \le \hat{a}_1 + t_{\alpha/2}\hat{\sigma}_{\hat{a}_1}, \\ &\hat{a}_2 - t_{\alpha/2}\hat{\sigma}_{\hat{a}_2} \le a_2 \le \hat{a}_2 + t_{\alpha/2}\hat{\sigma}_{\hat{a}_2}. \end{aligned} \tag{3.28}$$

Примечание – При использовании электронных таблиц Excel значение нижней и верхней границ интервалов 95%-го уровня доверия можно найти в таблице ВЫВОД ИТОГОВ пакета АНАЛИЗ ДАННЫХ – РЕГРЕССИЯ.

3.10 Прогнозирование по многофакторной линейной модели

Если эконометрическая модель адекватна по F-критерию Фишера, то ее можно использовать для экономического анализа, прогноза, принятия решения с фиксированной целью. Теория прогнозирования дает возможность получить точечные и интервальные прогнозные оценки.

Точечный прогноз индивидуального значения результативного показателя

При прогнозировании на основе временных рядов прогнозный период лежит после оцениваемого периода. При прогнозировании на основе пространственных рядов прогноз относится не к n элементам выборки, а к другим элементам, которые находятся за ее пределами, но относятся к той же генеральной совокупности.

Предположим, что известны прогнозные значения факторов для p-го периода или для p-го объекта, т. е. вектор $X_P = (1\ x_{p1}\ \ x_{p2}\ \ldots\ x_{pm})$, тогда можно получить прогнозное значение показателя по формуле

$$\hat{Y}_P = X_P \cdot \hat{A}. \tag{3.29}$$

При прогнозировании по двухфакторной модели прогнозное значение показателя определяется по формуле

$$\hat{y}_P = (1 x_{p_1} x_{p_2}) \cdot \begin{pmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} = \hat{a}_0 + \hat{a}_1 \cdot x_{p_1} + \hat{a}_2 \cdot x_{p_2}. \tag{3.30}$$

Интервальный прогноз математического ожидания результативного показателя

Для получения интервального прогноза математического ожидания показателя необходимо найти значение нижней и верхней границ прогнозного интервала по формулам:

$$\begin{aligned}&\textit{нижняя граница}:\ \hat{y}_P - t_{\alpha/2}\cdot\hat{\sigma}_{\bar{y}_p};\\&\textit{верхняя граница}:\ \hat{y}_P + t_{\alpha/2}\cdot\hat{\sigma}_{\bar{y}_p},\end{aligned}\tag{3.31}$$

где $\hat{y}_p$ – точечный прогноз индивидуального значения показателя;

$t_{\alpha/2}$ – табличное значение t-критерия Стьюдента с уровнем значимости α;

$\hat{\sigma}_{\bar{y}_p} = \sqrt{\hat{\sigma}^2_{\bar{y}_p}}$ – стандартная ошибка прогноза математического ожидания результативного показателя.

Оценка дисперсии прогноза математического ожидания рассчитывается по формуле

$$\hat{\sigma}^2_{\bar{y}_p} = X_p \hat{\Sigma}_{\hat{A}} X_P^T,\tag{3.32}$$

где $\hat{\Sigma}_{\hat{A}}$ – оценка ковариационной матрицы оценок параметров модели.

Исходя из этого, интервальный прогноз математического ожидания показателя для уровня значимости α будет иметь вид

$$\begin{aligned}&\hat{y}_p - t_{\alpha/2}\cdot\hat{\sigma}_e\sqrt{X_p(X^T\cdot X)^{-1}X_P^T} \le \bar{y}_p \le \hat{y}_p +\\&+ t_{\alpha/2}\cdot\hat{\sigma}_e\sqrt{X_p(X^T\cdot X)^{-1}X_P^T}.\end{aligned}\tag{3.33}$$

Интервальный прогноз индивидуального значения результативного показателя

Для получения интервального прогноза индивидуального значения показателя необходимо найти значение нижней и верхней границ прогнозного интервала по формулам:

$$\begin{aligned}&\textit{нижняя граница}:\ \hat{y}_P - t_{\alpha/2}\cdot\hat{\sigma}_{y_p};\\&\textit{верхняя граница}:\ \hat{y}_P + t_{\alpha/2}\cdot\hat{\sigma}_{y_p},\end{aligned}\tag{3.34}$$

где стандартная ошибка прогноза индивидуального значения рассчитывается по формуле

$$\hat{\sigma}_{y_p} = \sqrt{\hat{\sigma}^2_e + \hat{\sigma}^2_{\bar{y}_p}}\ .\tag{3.35}$$

Исходя из этого, интервальный прогноз индивидуального значения показателя для уровня значимости α будет иметь вид

$$\hat{y}_p - t_{\alpha/2} \cdot \hat{\sigma}_e \sqrt{1 + X_p (X^T \cdot X)^{-1} X_P^T} \le y_p \le \hat{y}_p + \\ + t_{\alpha/2} \cdot \hat{\sigma}_e \sqrt{1 + X_p (X^T \cdot X)^{-1} X_P^T}. \qquad (3.36)$$

Эконометрическая модель принятия решения с фиксированной целью дает возможность оценить, какое значение при равных других условиях должен принимать j-й фактор, чтобы при заданных значениях других факторов была получена необходимая величина результативного показателя (величина цели).

3.11 Общий алгоритм построения и анализа классической многофакторной линейной эконометрической модели

I Спецификация модели.

1 Идентификация переменных.

2 Оценка тесноты связи между показателем Y и факторами X_1 и X_2, а также между факторами.

а) Диаграмма рассеяния. Предварительные выводы о тесноте связи между показателями.

б) Коэффициенты парной корреляции, корреляционная матрица.

в) Коэффициенты частичной корреляции.

г) Выводы о существенности факторов и возможной мультиколлинеарности.

3 Общий вид линейной двухфакторной модели и оценка ее в матричной форме.

II Оценка параметров модели методом 1МНК в матричной форме.

III Коэффициент множественной детерминации и корреляции для оцененной модели.

1 Расчет коэффициентов множественной детерминации и корреляции.

2 Расчет коэффициента множественной детерминации.

3 Предварительные выводы об адекватности модели.

IV Оценка дисперсионно-ковариационной матрицы оценок параметров модели.

1 Оценка дисперсии остатков.

2 Расчет дисперсий и ковариаций оценок параметров модели.

3 Расчет стандартных ошибок оценок параметров модели.

4 Выводы о смещении оценок параметров модели.

V Проверка гипотез о статистической значимости оценок параметров модели на основе F- и t-критериев.

VI Построение доверительных интервалов для оценок параметров модели.

VII Прогнозирование на основе оцененной модели.

1 Точечный прогноз индивидуального значения показателя.

2 Доверительный интервал для прогноза математического ожидания показателя.

3 Доверительный интервал для прогноза индивидуального значения показателя.

VIII Экономический анализ по оцененной модели.

1 Средняя эффективность показателя.

2 Предельная эффективность показателя.

3 Частичные коэффициенты эластичности и общая эластичность.

4 Предельная норма замещения факторов.

3.12 Реализация линейной многофакторной эконометрической модели средствами табличного процессора MS Excel

По статистическим данным 18-ти предприятий производственного объединения построить эконометрическую модель, которая характеризует зависимость между общими затратами, объемом грузооборота и фондоемкостью предприятия, оценить ее адекватность и сделать прогноз затрат одного из предприятий объединения при соответствующем изменении объема грузооборота и фондоемкости. Сделать экономический анализ по модели. Исходные данные приведены на рисунке 3.1.

Решение задачи проведем в среде Excel по указанному выше алгоритму.

I Спецификация модели.

1 Идентификация переменных:

Y – общие затраты – результирующий показатель;

X_1 – грузооборот – показатель-фактор;

X_2 – фондоемкость – показатель-фактор.

Блок данных формируется в массиве C3:E18, в ячейки B3:B18 вводятся номера наблюдений (см. рисунок 3.1).

№	y	x1	x2	y*x1	y*x2	x1*x2	y*y	x1*x1	x2*x2
1	20	45	1,5	900	30	67,5	400	2025	2,25
2	32	75	1,6	2400	51,2	120	1024	5625	2,56
3	48	125	1,9	6000	91,2	237,5	2304	15625	3,61
4	65	223	1,8	14495	117	401,4	4225	49729	3,24
5	45	92	3,4	4140	153	312,8	2025	8464	11,56
6	64	146	3,6	9344	230,4	525,6	4096	21316	12,96
7	79	227	3,5	17933	276,5	794,5	6241	51529	12,25
8	104	358	5,5	37232	572	1969	10816	128164	30,25
9	68	135	5,4	9180	367,2	729	4624	18225	29,16
10	93	218	5,4	20274	502,2	1177,2	8649	47524	29,16
11	117	331	5,3	38727	620,1	1754,3	13689	109561	28,09
12	145	490	8,5	71050	1232,5	4165	21025	240100	72,25
13	91	175	8,3	15925	755,3	1452,5	8281	30625	68,89
14	131	205	8,1	26855	1061,1	1660,5	17161	42025	65,61
15	167	468	7,3	78156	1219,1	3416,4	27889	219024	53,29
16	195	749	8,4	146055	1638	6291,6	38025	561001	70,56
сумма	1464	4062	79,5	498666	8916,8	25074,8	170474	1550562	495,69
ср	91,5	253,875	4,96875	31166,63	557,3	1567,175	10654,6	96910,1	30,9806

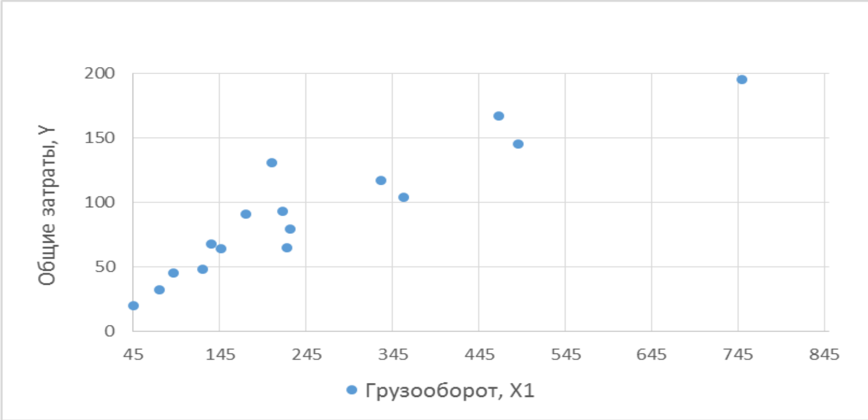
Рисунок 3.1 – Исходные данные и элементарные преобразования этих данных для оценки модели

2 Оценка тесноты связи между показателем y и факторами x_1 и x_2, а также между факторами.

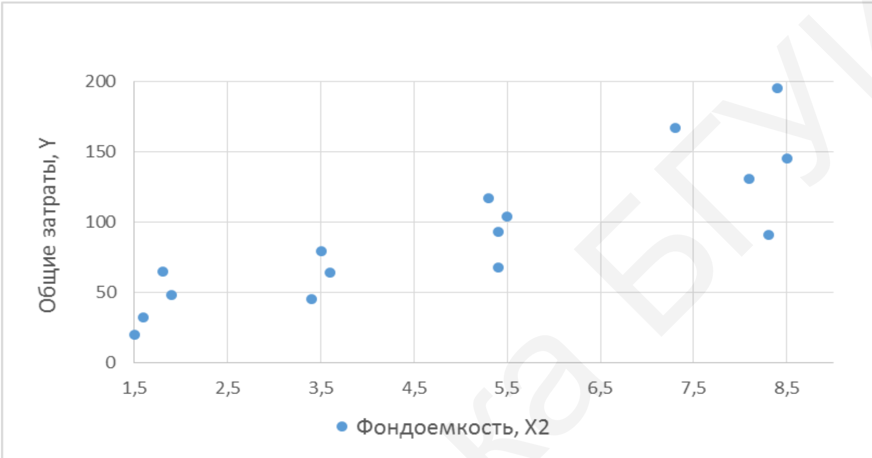
а) Диаграмма рассеяния. Предварительные выводы о тесноте связи между показателями.

В модель необходимо в первую очередь включить факторы, которые коррелируют с показателем и не коррелируют между собой, т. е. не являются мультиколлинеарными. Для предварительного анализа тесноты связи и типа аналитической зависимости между исследуемыми показателями целесообразно построить соответствующие диаграммы рассеяния.

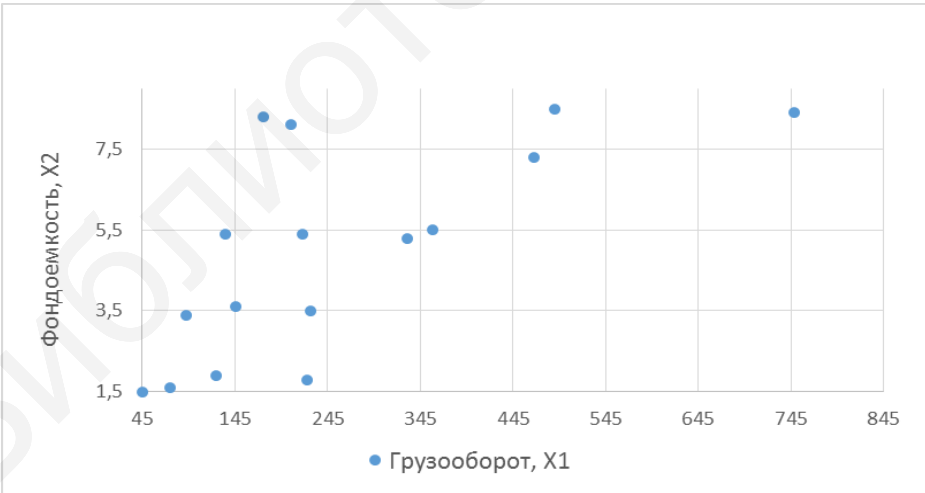
На основе визуального исследования диаграмм рассеяния можно сделать предварительные выводы о тесноте связи между показателем y и факторами x_1 и x_2, а также между факторами (рисунок 3.2).



а



б



в

а – $r_{yx_1} \approx 0{,}9$ – прямая тесная линейная зависимость;

б – $r_{yx_{21}} \approx 0{,}8$ – прямая тесная линейная корреляционная зависимость;

в – $r_{x_1x_2} \approx 0{,}7$ – прямая средняя линейная корреляционная зависимость

Рисунок 3.2 – Диаграмма рассеяния

Таким образом, грузооборот и фондоемкость оказывают существенное влияние на показатель общих затрат (рисунок 3.2, а, б), поэтому эти показатели являются значимыми, их целесообразно ввести в модель. Но между самими факторами имеет место средняя корреляционная зависимость (рисунок 3.2, в), что может свидетельствовать о возможной мультиколлинеарности и ее негативных последствиях.

б) Коэффициенты парной корреляции, корреляционная матрица.

Для количественной оценки тесноты связи между показателем y и факторами x_1 и x_2, а также между факторами рассчитываются парные коэффициенты корреляции по формулам (3.8), а затем составляют корреляционную матрицу (3.7) исходя из ее свойств.

Расчеты и проверка при помощи встроенных статистических функций приведены на рисунках 3.3, 3.4.

	A	B	C	D	E	F	G	H	I	J	K
1											
2		№	y	x1	x2	y*x1	y*x2	x1*x2	y*y	x1*x1	x2*x2
3		1	20	45	1,5	=C3*D3	=C3*E3	=D3*E3	=C3*C3	=D3*D3	=E3*E3
4		2	32	75	1,6	=C4*D4	=C4*E4	=D4*E4	=C4*C4	=D4*D4	=E4*E4
5		3	48	125	1,9	=C5*D5	=C5*E5	=D5*E5	=C5*C5	=D5*D5	=E5*E5
6		4	65	223	1,8	=C6*D6	=C6*E6	=D6*E6	=C6*C6	=D6*D6	=E6*E6
7		5	45	92	3,4	=C7*D7	=C7*E7	=D7*E7	=C7*C7	=D7*D7	=E7*E7
8		6	64	146	3,6	=C8*D8	=C8*E8	=D8*E8	=C8*C8	=D8*D8	=E8*E8
9		7	79	227	3,5	=C9*D9	=C9*E9	=D9*E9	=C9*C9	=D9*D9	=E9*E9
10		8	104	358	5,5	=C10*D10	=C10*E10	=D10*E10	=C10*C10	=D10*D10	=E10*E10
11		9	68	135	5,4	=C11*D11	=C11*E11	=D11*E11	=C11*C11	=D11*D11	=E11*E11
12		10	93	218	5,4	=C12*D12	=C12*E12	=D12*E12	=C12*C12	=D12*D12	=E12*E12
13		11	117	331	5,3	=C13*D13	=C13*E13	=D13*E13	=C13*C13	=D13*D13	=E13*E13
14		12	145	490	8,5	=C14*D14	=C14*E14	=D14*E14	=C14*C14	=D14*D14	=E14*E14
15		13	91	175	8,3	=C15*D15	=C15*E15	=D15*E15	=C15*C15	=D15*D15	=E15*E15
16		14	131	205	8,1	=C16*D16	=C16*E16	=D16*E16	=C16*C16	=D16*D16	=E16*E16
17		15	167	468	7,3	=C17*D17	=C17*E17	=D17*E17	=C17*C17	=D17*D17	=E17*E17
18		16	195	749	8,4	=C18*D18	=C18*E18	=D18*E18	=C18*C18	=D18*D18	=E18*E18
19		сумма	=СУММ(C3:C18)	=СУММ(D3:D18)	=СУММ(E3:E18)	=СУММ(F3:F18)	=СУММ(G3:G18)	=СУММ(H3:H18)	=СУММ(I3:I18)	=СУММ(J3:J18)	=СУММ(K3:K18)
20		ср	=СРЗНАЧ(C3:C18)	=СРЗНАЧ(D3:D18)	=СРЗНАЧ(E3:E18)	=СРЗНАЧ(F3:F18)	=СРЗНАЧ(G3:G18)	=СРЗНАЧ(H3:H18)	=СРЗНАЧ(I3:I18)	=СРЗНАЧ(J3:J18)	=СРЗНАЧ(K3:K18)
21											

Рисунок 3.3 – Исходные данные и элементарные преобразования этих данных для оценки модели (формулы)

	A	B	C	D	E
22					
23			По формулам	Мастер функций	
24		Дисп y	2282,375	2282,375	
25		Дисп x1	32457,60938	32457,6094	
26		Дисп x2	6,292148438	6,29214844	
27		СКО y	47,77420852	47,7742085	
28		СКО x1	180,159955	180,159955	
29		СКО x2	2,508415523	2,50841552	
30		ковариация y x1	7937,0625	7937,0625	
31		ковариация y x2	102,659375	102,659375	
32		ковариация x1x2	305,7335938	305,733594	
33		коэффициенты парной корреляции			
34		ryx1	0,922163726	0,92216373	
35		ryx2	0,856654371	0,85665437	
36		rx1x2	0,676527441	0,67652744	
37					

а

	A	B	C	D
22				
23			По формулам	Мастер функций
24		Дисп y	=I20-C20*C20	=ДИСПР(C3:C18)
25		Дисп x1	=J20-D20*D20	=ДИСПР(D3:D18)
26		Дисп x2	=K20-E20*E20	=ДИСПР(E3:E18)
27		СКО y	=C24^0,5	=СТАНДОТКЛОНП(C3:C18)
28		СКО x1	=C25^0,5	=КОРЕНЬ(D25)
29		СКО x2	=C26^0,5	=КОРЕНЬ(D26)
30		ковариация y x1	=F20-C20*D20	=КОВАР(C3:C18;D3:D18)
31		ковариация y x2	=G20-C20*E20	=КОВАР(C3:C18;E3:E18)
32		ковариация x1x2	=H20-D20*E20	=КОВАР(D3:D18;E3:E18)
33		коэффициенты пар		
34		ryx1	=C30/C28/C27	=КОРРЕЛ(C3:C18;D3:D18)
35		ryx2	=C31/C29/C27	=КОРРЕЛ(C3:C18;E3:E18)
36		rx1x2	=C32/C28/C29	=КОРРЕЛ(D3:D18;E3:E18)
37				

б

а – значения; б – формулы

Рисунок 3.4 – Расчет коэффициентов парной корреляции в) Коэффициенты частичной корреляции.

В многофакторной модели коэффициенты парной корреляции отображают нечистую связь между факторами и показателем. Поэтому при построении данной модели целесообразно оценить связь между показателем и одним фактором при условии, что влияние другого фактора не учитывается. Для вычисления такой чистой связи рассчитывают коэффициенты частичной корреляции по формулам (3.10). Расчет представлен на рисунках 3.5 и 3.6.

	A	B	C	D	E	F
41						
42		* по определению				
43		ryx1(x2)	0,901825898			
44		ryx2(x1)	0,817231878			
45		rx1x2(y)	-0,568530062			
46		корреляционная матрица				
47			y	x1	x2	
48		y	1	0,92216373	0,856654371	
49		x1	0,922163726	1	0,676527441	
50		x2	0,856654371	0,67652744	1	
51						

Рисунок 3.5 – Расчет коэффициентов частичной корреляции (значение)

	A	B	C	D	E
41					
42		* по определению			
43		ryx1(x2)	=(C34-C35*C36)/((1-C35*C35)*(1-C36*C36))^0,5		
44		ryx2(x1)	=(C35-C34*C36)/((1-C34*C34)*(1-D36*D36))^0,5		
45		rx1x2(y)	=(D36-D34*D35)/((1-D34*D34)*(1-D35*D35))^0,5		
46		корреляционная ма			
47			y	x1	x2
48		y	1	=C34	=C35
49		x1	=C34	1	=C36
50		x2	=C35	=C36	1
51					

Рисунок 3.6 – Расчет коэффициентов частичной корреляции (формулы)

г) Выводы о существенности факторов и возможной мультиколлинеарности.

При помощи полученных коэффициентов парной и частичной корреляции можно сделать выводы о существенности факторов и проверить их на мультиколлинеарность – линейную зависимость или сильную корреляцию:

- поскольку коэффициент парной корреляции между затратами оборота и грузооборота $r_{yx_1} = 0{,}9222$ и соответствующий коэффициент частичной корреляции $r_{yx_1(x_2)} = 0{,}9018$, то это свидетельствует о том, что грузооборот существенно влияет на затраты оборота и его целесообразно ввести в модель;

- поскольку коэффициент парной корреляции между затратами оборота и фондоемкостью $r_{yx_2} = 0{,}8567$ и соответствующий коэффициент частичной корреляции $r_{yx2(x_1)} = 08172$, то это свидетельствует о том, что между затратами оборота и фондоемкостью также существует сильная связь и этот фактор целесообразно ввести в модель;

- поскольку коэффициент парной корреляции между факторами $r_{x_1x_2} = 0{,}6765$, то это свидетельствует о том, что между факторами существует средняя, близкая к сильной, корреляционная зависимость. Чистая связь между факторами отрицательная $r_{x_1x_2(y)} = -0{,}5685$ и также средняя, следовательно, можно сделать вывод о возможной неэкстремальной мультиколлинеарности факторов.

3 Общий вид линейной двухфакторной модели и оценка ее в матричной форме.

Допустим, что между x_1 и x_2 и показателем y линейная стохастическая зависимость, которая может быть описана линейной двухфаторной моделью в виде (3.1), т. е. $y = a_0 + a_1x_1 + a_2x_2 + e$, ее оценкой будет уравнение $\hat{y} = \hat{a}_0 + \hat{a}_1x_1 + \hat{a}_2x_2$.

В матричной форме модель и ее оценка могут быть записаны в виде (3.3) и (3.4) соответственно. Для имеющихся статистических данных соответствующие матрицы будут иметь вид

$$y = \begin{pmatrix} 20 \\ 32 \\ 48 \\ 65 \\ 45 \\ 64 \\ 79 \\ 104 \\ 68 \\ 93 \\ 117 \\ 145 \\ 91 \\ 131 \\ 167 \\ 195 \end{pmatrix}, \quad x = \begin{pmatrix} 1 & 45 & 1{,}5 \\ 1 & 75 & 1{,}6 \\ 1 & 125 & 1{,}9 \\ 1 & 223 & 1{,}8 \\ 1 & 92 & 3{,}4 \\ 1 & 146 & 3{,}6 \\ 1 & 227 & 3{,}5 \\ 1 & 358 & 3{,}5 \\ 1 & 135 & 5{,}4 \\ 1 & 218 & 5{,}4 \\ 1 & 331 & 5{,}3 \\ 1 & 490 & 8{,}5 \\ 1 & 175 & 8{,}3 \\ 1 & 205 & 8{,}1 \\ 1 & 468 & 7{,}3 \\ 1 & 749 & 8{,}4 \end{pmatrix}, \quad A = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}, \quad \hat{A} = \begin{pmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \end{pmatrix}.$$

II Оценка параметров модели методом 1МНК в матричной форме.

Алгоритм вычисления параметров модели:

а) Вычислить матрицу моментов $X^T \cdot X$, используя встроенную функцию МУМНОЖ (рисунки 3.7 и 3.8).

	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN
1																							
2																		1	45	1,5			
3																		1	75	1,6			
4																		1	125	1,9			
5																		1	223	1,8			
6																		1	92	3,4			
7																		1	146	3,6			
8		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	227	3,5	16	4062	79,5
9		45	75	125	223	92	146	227	358	135	218	331	490	175	205	468	749	1	358	5,5	4062	1550562	25074,8
10		1,5	1,6	1,9	1,8	3,4	3,6	3,5	5,5	5,4	5,4	5,3	8,5	8,3	8,1	7,3	8,4	1	135	5,4	79,5	25074,8	495,69
11																		1	218	5,4			
12																		1	331	5,3			
13																		1	490	8,5			
14																		1	175	8,3			
15																		1	205	8,1			
16																		1	468	7,3			
17																		1	749	8,4			
18																							

Рисунок 3.7 – Матрица моментов $X^T \cdot X$ (значения)

	AF	AG	AH	AI	AJ	AK	AL	AM
1								
2				1	45	1,5		
3				1	75	1,6		
4				1	125	1,9		
5				1	223	1,8		
6				1	92	3,4		
7				1	146	3,6		
8	=ТРАНСП(AI2:AK17)	=ТРАНСП(AI2:AK17)	=ТРАНСП(AI2:AK17)	1	227	3,5	=МУМНОЖ(S8:AH10;AI2:AK17)	=МУМНОЖ(S8:AH1
9	=ТРАНСП(AI2:AK17)	=ТРАНСП(AI2:AK17)	=ТРАНСП(AI2:AK17)	1	358	5,5	=МУМНОЖ(S8:AH10;AI2:AK17)	=МУМНОЖ(S8:AH1
10	=ТРАНСП(AI2:AK17)	=ТРАНСП(AI2:AK17)	=ТРАНСП(AI2:AK17)	1	135	5,4	=МУМНОЖ(S8:AH10;AI2:AK17)	=МУМНОЖ(S8:AH1
11				1	218	5,4		
12				1	331	5,3		
13				1	490	8,5		
14				1	175	8,3		
15				1	205	8,1		
16				1	468	7,3		
17				1	749	8,4		
18								

Рисунок 3.8 – Матрица моментов $X^T \cdot X$ (формулы)

б) Вычисляем матрицу ошибок $(X^T \cdot X)^{-1}$, используя функцию МОБР (рисунки 3.9 и 3.10).

	R	S	T	U	V
17					
18					
19		0,30828	-4,42E-05	-0,04721	
20		-4E-05	3,55E-06	-0,00017	
21		-0,0472	-0,000173	0,018316	
22					
23					

Рисунок 3.9 – Матрица ошибок $(X^T \cdot X)^{-1}$ (значения)

	R	S	T	U
18				
19		=МОБР(AL8:AN10)	=МОБР(AL8:AN10)	=МОБР(AL8:AN10)
20		=МОБР(AL8:AN10)	=МОБР(AL8:AN10)	=МОБР(AL8:AN10)
21		=МОБР(AL8:AN10)	=МОБР(AL8:AN10)	=МОБР(AL8:AN10)
22				

Рисунок 3.10 – Матрица ошибок $(X^T \cdot X)^{-1}$ (формулы)

в) Вычисляем матрицу $X^T \cdot Y$, используя функцию МУМНОЖ (рисунки 3.11 и 3.12).

	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF	BG	BH
1																				
2																		20		
3																		32		
4																		48		
5																		65		
6																		45		
7																		64		
8		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	79	1464	
9		45	75	125	223	92	146	227	358	135	218	331	490	175	205	468	749	104	498666	
10		1,5	1,6	1,9	1,8	3,4	3,6	3,5	5,5	5,4	5,4	5,3	8,5	8,3	8,1	7,3	8,4	68	8916,8	
11																		93		
12																		117		
13																		145		
14																		91		
15																		131		
16																		167		
17																		195		
18																				

Рисунок 3.11 – Матрица $X^T \cdot Y$ (значения)

	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF	BG
1																			
2																		20	
3																		32	
4																		48	
5																		65	
6																		45	
7																		64	
8		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	79	=МУМНОЖ(AP8:BE10;BF2:BF17)
9		45	75	125	223	92	146	227	358	135	218	331	490	175	205	468	749	104	=МУМНОЖ(AP8:BE10;BF2:BF17)
10		1,5	1,6	1,9	1,8	3,4	3,6	3,5	5,5	5,4	5,4	5,3	8,5	8,3	8,1	7,3	8,4	68	=МУМНОЖ(AP8:BE10;BF2:BF17)
11																		93	
12																		117	
13																		145	
14																		91	
15																		131	
16																		167	
17																		195	
18																			
19																			

Рисунок 3.12 – Матрица $X^T \cdot Y$ (формулы)

г) Вычисляем вектор оценок параметров модели как произведение матриц $(X^T \cdot X)^{-1}$ и $X^T \cdot Y$ (рисунки 3.13 и 3.14).

д) Для проверки правильности определения оценок параметров модели регрессионных статистик в Excel воспользуемся пакетом Анализ данных (рисунок 3.15) для сравнения полученных значений оценок параметров модели по формулам и при помощи Пакета анализа скопируем с листа «Регрессия» значения ячеек столбца Коэффициенты строк Y-пересечение, X_1 и X_2. Полученные значения должны совпасть.

	V	W	X	Y	Z	AA
19		по формуле		Регрессия		
20				коэффициенты		
21		8,3476	a0	Y-пересечение	8,3476	
22		0,16753	a1	Переменная X	0,16753	
23		8,17526	a2	Переменная X	8,17526	
24						

Рисунок 3.13 – Вектор оценок параметров модели (значения)

	W	X	Y	Z	AA
19	по формуле		Регрессия		
20			коэффициенты		
21	=МУМНОЖ(S19:U21;BG8:BG10)	a0	=Регрессия!A17	=Регрессия!B17	
22	=МУМНОЖ(S19:U21;BG8:BG10)	a1	=Регрессия!A18	=Регрессия!B18	
23	=МУМНОЖ(S19:U21;BG8:BG10)	a2	=Регрессия!A19	=Регрессия!B19	

Рисунок 3.14 – Вектор оценок параметров модели (формулы)

	A	B	C	D	E	F	G	H	I
1	ВЫВОД ИТОГОВ								
2									
3	*Регрессионная статистика*								
4	Множественный R	0,974838							
5	R-квадрат	0,950308							
6	Нормированный R-квадрат	0,942664							
7	Стандартная ошибка	11,81472							
8	Наблюдения	16							
9									
10	Дисперсионный анализ								
11		*df*	*SS*	*MS*	*F*	*Значимость F*			
12	Регрессия	2	34703,36	17351,68	124,3067	3,35615E-09			
13	Остаток	13	1814,639	139,5876					
14	Итого	15	36518						
15									
16		*Коэффи-циенты*	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*	*Нижние 95,0%*	*Верхние 95,0%*
17	Y-пересечение	8,3476	6,559893	1,272521	0,225477	-5,824186575	22,51939	-5,82419	22,51939
18	Переменная X 1	0,16753	0,022263	7,52507	4,34E-06	0,119433633	0,215626	0,119434	0,215626
19	Переменная X 2	8,175258	1,598968	5,112833	0,000199	4,720896785	11,62962	4,720897	11,62962
20									

Рисунок 3.15 – Итоговый лист

Таким образом, оцененная эконометрическая модель имеет вид $\hat{y} = 8{,}3476 + 0{,}16753x_1 + 8{,}175285x_2$.

III Коэффициент множественной детерминации и корреляции для оцененной модели.

1 Расчет коэффициентов множественной детерминации и корреляции.

Для оценки степени соответствия полученной модели наблюдаемым значениям, т. е. предварительной оценки адекватности модели, вычислим коэффициенты множественной детерминации и корреляции.

Алгоритм расчета коэффициентов множественной детерминации и корреляции.

а) Скопируем с итогового листа РЕГРЕССИЯ значения столбцов «Предсказанное *Y*» и «Остатки» (рисунки 3.16 и 3.17).

б) Рассчитаем средние значения для фактических и оцененных данных показателя *y*, используя функцию СРЗНАЧ.

в) В ячейку F57 введем формулу для общих отклонений $y_i - \bar{y}$ и скопируем ее в остальные ячейки этого столбца.

г) В ячейках E74 и F74 вычислим суммы квадратов общих отклонений и отклонений, которые не объясняются регрессией (остатков), используя встроенную функцию СУММКВ.

д) В ячейку H69 введем формулу расчета коэффициента множественной детерминации $R^2 = 1 - \dfrac{\sum_{i=1}^{n} e^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$.

е) В ячейку H63 введем формулу для расчета коэффициента множественной корреляции $R = \sqrt{R^2}$.

ж) Для проверки полученных коэффициентов скопируем с итогового листа РЕГРЕССИЯ значения ячеек «*R*-квадрат» и «Множественный *R*». Можно заметить, что значения совпали.

	A	B	C	D	E	F	G	H	I	J
55										
56		№	факт	предсказ	Остатки	yi-_yi				
57		1	20	28,1493209	-8,14932087	-71,5				
58		2	32	33,9927357	-1,99273573	-59,5				
59		3	48	44,8217949	3,178205076	-43,5				
60		4	65	60,4221734	4,577826645	-26,5				
61		5	45	51,5562041	-6,55620414	-46,5				
62		6	64	62,2378561	1,762143938	-27,5				
63		7	79	74,9902307	4,009769305	-12,5				
64		8	104	113,287129	-9,28712913	12,5				
65		9	68	75,1104947	-7,11049469	-23,5				
66		10	93	89,0154544	3,984545597	1,5				
67		11	117	107,128777	9,871222639	25,5		по формуле		регрессия
68		12	145	159,926815	-14,9268153	53,5		коэф.детерминации		
69		13	91	105,519929	-14,5199286	-0,5		0,95030837	р-квадрат	0,950308
70		14	131	108,910766	22,08923394	39,5				
71		15	167	146,430854	20,56914639	75,5				
72		16	195	202,499465	-7,49946501	103,5			коэф.мн.корреляции	
73		срзнач	91,5	91,5				0,97483761	множ р	0,974838
74		суммкв			1814,638824	36518				

Рисунок 3.16 – Расчет коэффициентов R и R^2 (значения)

	A	B	C	D	E	F
55						
56		№	факт	предсказ	Остатки	yi-_yi
57		1	=C3	=Регрессия!B26	=Регрессия!C26	=C57-$D$73
58		2	=C4	=Регрессия!B27	=Регрессия!C27	=C58-$D$73
59		3	=C5	=Регрессия!B28	=Регрессия!C28	=C59-$D$73
60		4	=C6	=Регрессия!B29	=Регрессия!C29	=C60-$D$73
61		5	=C7	=Регрессия!B30	=Регрессия!C30	=C61-$D$73
62		6	=C8	=Регрессия!B31	=Регрессия!C31	=C62-$D$73
63		7	=C9	=Регрессия!B32	=Регрессия!C32	=C63-$D$73
64		8	=C10	=Регрессия!B33	=Регрессия!C33	=C64-$D$73
65		9	=C11	=Регрессия!B34	=Регрессия!C34	=C65-$D$73
66		10	=C12	=Регрессия!B35	=Регрессия!C35	=C66-$D$73
67		11	=C13	=Регрессия!B36	=Регрессия!C36	=C67-$D$73
68		12	=C14	=Регрессия!B37	=Регрессия!C37	=C68-$D$73
69		13	=C15	=Регрессия!B38	=Регрессия!C38	=C69-$D$73
70		14	=C16	=Регрессия!B39	=Регрессия!C39	=C70-$D$73
71		15	=C17	=Регрессия!B40	=Регрессия!C40	=C71-$D$73
72		16	=C18	=Регрессия!B41	=Регрессия!C41	=C72-$D$73
73		срзнач	=СРЗНАЧ(C57:C72)	=СРЗНАЧ(D57:D72)		
74		суммкв			=СУММКВ(E57:E72)	=СУММКВ(F57:F72)
75						

а

	H	I	J
66			
67	по формуле		регрессия
68	коэф.детерминации		
69	=1-E74/F74	р-квадрат	=Регрессия!B5
70			
71			
72		коэф.мн.корреляции	
73	=H69^0,5	множ р	=Регрессия!B4
74			
64			

б

а – таблица расчетов; б – расчет R и R^2

Рисунок 3.17 – Расчет коэффициентов R и R^2 (формулы)

2 Расчет коэффициента множественной детерминации.

Рассчитаем коэффициент детерминации по формуле (3.13) в ячейках C76 и C77 (рисунок 3.18).

	A	B	C	D
76		d1^2	0,582591899	
77		d2^2	0,367716475	
78		r^2	0,950308373	

а

	A	B	C
75			
76		d1^2	=C34*Z22*D28/D27
77		d2^2	=C35*Z23*D29/D27
78		r^2	=СУММ(C76:C77)

б

а – значения; б – формулы

Рисунок 3.18 – Расчет коэффициентов детерминации

3 Предварительные выводы об адекватности модели.

При помощи полученных коэффициентов множественной детерминации, корреляции и отдельной детерминации можно сделать предварительные выводы об адекватности модели:

- поскольку коэффициент множественной детерминации $R^2 = 0{,}9503$, то это свидетельствует о том, что вариация общих затрат на предприятиях объединения на 95,03 % определяется вариацией грузооборота и фондоемкости и на 4,97 % вариацией факторов, которые не включены в модель;

- поскольку коэффициент отдельной детерминации $d_1^2 = 0{,}5826$, $d_2^2 = 0{,}3677$, то это свидетельствует о том, что вариация общих затрат на предприятиях объединения на 58,26 % определяется вариацией грузооборота и на 36,77 % вариацией фондоемкости;

- коэффициент множественной корреляции $R = 0{,}9748$ характеризует достаточно тесную связь между общими затратами и факторами, которые их обуславливают.

Полученную модель можно считать адекватной статистическим данным.

IV Оценка дисперсионно-ковариационной матрицы оценок параметров модели.

1 Оценка дисперсии остатков.

Вычислим оценку дисперсии отклонений по формуле $\hat{\sigma}_e^2 = \dfrac{\sum\limits_{i=1}^{n} {e_i}^2}{n-m-1}$, которую введем в ячейку B83 (рисунок 3.19).

	A	B	C	D
81				
82		формула	регрессия	MS
83		139,588	остаток	139,587602

а

	A	B	C	D
81				
82		формула	регрессия	MS
83		=E74/13	остаток	=Регрессия!D13

б

а – значения; б – формулы

Рисунок 3.19 – Оценка дисперсии остатков

2 Расчет дисперсий и ковариаций оценок параметров модели.

Для получения оценок ковариаций и дисперсий оценок параметров модели необходимо составить ковариационную матрицу по формуле (3.19), т. е.

$$\hat{\Sigma}_{\hat{A}} = \hat{\sigma}_e^2 \cdot \left(X^T \cdot X\right)^{-1} = \begin{pmatrix} \hat{\sigma}_{\hat{a}_0}^2 & \hat{\sigma}_{\hat{a}_0\hat{a}_1} & \hat{\sigma}_{\hat{a}_0\hat{a}_2} \\ \hat{\sigma}_{\hat{a}_1\hat{a}_0} & \hat{\sigma}_{\hat{a}_1}^2 & \hat{\sigma}_{\hat{a}_1\hat{a}_2} \\ \hat{\sigma}_{\hat{a}_2\hat{a}_0} & \hat{\sigma}_{\hat{a}_2\hat{a}_1} & \hat{\sigma}_{\hat{a}_2}^2 \end{pmatrix}.$$

Введем в ячейку B87 (рисунок 3.20) значения дисперсии отклонений из ячейки B83, а в диапазон C86:E88 – элементы матрицы ошибок (см. рисунок 3.9). Элементы ковариационной матрицы рассчитаем как произведение скалярной величины (дисперсии) и матрицы ошибок. В ячейку F86 (рисунок 3.20) вводим формулу произведения на первый элемент матрицы ошибок, используя абсолютные и относительные адреса. Далее последовательным копированием по строкам и столбцам, изменяя номер строки, получаем ковариационную матрицу в диапазоне F86:H88. Таким образом, получили дисперсии оценок параметров модели: $\hat{\sigma}_{\hat{a}0}^2$ =43,03, $\hat{\sigma}_{\hat{a}1}^2$ =0,0005, $\hat{\sigma}_{\hat{a}2}^2$=2,5567, стоящих на главной диагонали матрицы.

	A	B	C	D	E	F	G	H
85								
86			0,308280911	-4,4188E-05	-0,04720756	43,032193	-0,00616815	-6,5895907
87		139,588	-4,41884E-05	3,5507E-06	-0,00017253	-0,0061681	0,000495635	-0,0240828
88			-0,047207564	-0,00017253	0,018316098	-6,5895907	-0,024082769	2,55670013
89								

а

	A	B	C	D	E	F	G	H
85								
86			=S19	=T19	=U19	=C86*$B$87	=D86*$B$87	=E86*$B$87
87		=B83	=S20	=T20	=U20	=C87*$B$87	=D87*$B$87	=E87*$B$87
88			=S21	=T21	=U21	=C88*$B$87	=D88*$B$87	=E88*$B$87
89								

б

а – значения; б – формулы

Рисунок 3.20 – Оценка ковариационной матрицы оценок параметров модели

3 Расчет стандартных ошибок оценок параметров модели.

Для получения стандартной ошибки оценки параметров модели необходимо воспользоваться формулами $\hat{\sigma}_{\hat{a}_0} = \sqrt{\hat{\sigma}^2_{\hat{a}_0}}$, $\hat{\sigma}_{\hat{a}_1} = \sqrt{\hat{\sigma}^2_{\hat{a}_1}}$, $\hat{\sigma}_{\hat{a}_2} = \sqrt{\hat{\sigma}^2_{\hat{a}_2}}$ (рисунки 3.21 и 3.22).

	A	B	C	D	E	F	G
89							
90		формула	регрессия				
91			Стандартная ошибка		выводы о смещенность оценок параметров		
92		6,55989	6,559892761		78,584176	оценка смещена	
93		0,02226	0,022262868		13,28891334	оценка смещена	
94		1,59897	1,598968458		19,55862965	оценка смещена	

Рисунок 3.21 – Расчет стандартных ошибок оценок параметров модели. Выводы о смещении оценок параметров модели (значения)

	A	B	C	D	E	F
91			=Регрессия!C16		выводы о смеще	
92		=F86^0,5	=Регрессия!C17		=(B92/W21)*100	=ЕСЛИ(E92>10; "оценка смещена";"оценка не смещена")
93		=G87^0,5	=Регрессия!C18		=(B93/W22)*100	=ЕСЛИ(E93>10; "оценка смещена";"оценка не смещена")
94		=H88^0,5	=Регрессия!C19		=(B94/W23)*100	=ЕСЛИ(E94>10; "оценка смещена";"оценка не смещена")

Рисунок 3.22 – Расчет стандартных ошибок оценок параметров модели. Выводы о смещении оценок параметров модели (формулы)

4 Выводы о смещении оценок параметров модели.

Сравним каждую стандартную ошибку с соответствующим значением оценки параметра при помощи формул (3.21). Выводы о смещении (несмещении) оценок параметров модели сделаем, используя встроенную логическую функцию ЕСЛИ. Для всех параметров стандартные ошибки превышают 10 % от абсолютных значений параметров, что свидетельствует о том, что все оценки параметров модели смещены.

V Проверка гипотез о статистической значимости оценок параметров модели на основе *F*- и *t*-критериев.

1 Проверка адекватности модели по критерию Фишера.

Проверку адекватности модели по критерию Фишера проведем в соответствии с приведенным ниже алгоритмом.

Шаг 1. Формирование нулевой и альтернативной гипотез.

H_0: $\hat{a}_0 = \hat{a}_1 = \hat{a}_2 = 0$, т. е. все оценки модели незначимы, модель неадекватна.

H_1: Хотя бы одно значение $\hat{a}_j$ отлично от нуля, т. е. $\hat{a}_0 \neq 0$, $j = \overline{0;\ 2}$.

Шаг 2. Выбор соответствующего уровня значимости.

Выбираем уровень значимости α = 0,05, т. е. доверительная вероятность P = 0,95.

Шаг 3. Вычисляем расчетное значение *F*-критерия.

Для вычисления *F*-критерия в ячейку B99 введем формулу (3.24) (рисунки 3.23 и 3.24). Для проверки скопируем значение с итогового листа РЕГРЕССИЯ значение *F*-критерия. Значения совпали. $F_{расч}$ = 124,307.

Шаг 4. Определение по статистическим таблицам *F*-распределения Фишера и критического значения *F*-критерия.

Критическое значение *F*-критерия находим по статистическим таблицам *F*-распределения Фишера по соответствующим значениям:

- доверительная вероятность P = 0,95;
- степени свободы k_1 = 2, k_2 = 13.

Полученное значение критерия $F_{кр}$ = 3,81 вводим в ячейку или используем встроенную статистическую функцию FРАСПОБР.

Шаг 5. Сравнение расчетного значения *F*-критерия с критическим и интерпретация результатов.

Вывод о принятии нулевой гипотезы, т. е. об адекватности модели делаем в ячейке F99, используя встроенную функцию ЕСЛИ.

Поскольку $F_{расч} > F_{кр}$, то нулевая гипотеза о незначимости факторов отклоняется с риском ошибиться не более чем в 5 % случаев, т. е. с надежностью P = 0,95 можно сказать, что принятая модель адекватна статистическим данным и на основе этой модели можно делать экономический анализ и прогнозирование.

	A	B	C	D	E	F
95						
96		критерий фишера				
97		по форм	регрессия		р 0,95	
98			F		а табл(2,13)	3,18
99		124,307	124,3067461		модель адекватна	
100		критерий стьюдента				
101		по форм	регрессия		P 0,95	
102			т-статистика		табл(13)	2,16
103		1,27252	1,272520819	a0	параметр не явл. значимым	
104		7,52507	7,525069768	a1	параметр значим	
105		5,11283	5,112832637	a2	параметр значим	
106						

Рисунок 3.23 – Проверка гипотез о статистической значимости оценок параметров модели на основе *F*- и *t*-критериев (значения)

	C	D	E	F
95				
96	критерий фишера			
97	регрессия		р 0,95	
98	F		а табл(2,13)	3,18
99	=Регрессия!E12		=ЕСЛИ(B99>F98;"модель адекватна";"модель неадекватна")	
100	критерий стьюдента			
101	регрессия		P 0,95	
102	т-статистика		табл(13)	2,16
103	=Регрессия!D17	a0	=ЕСЛИ($B103>$F$102;"параметр значим ";"параметр незначим ")	
104	=Регрессия!D18	a1	=ЕСЛИ($B104>$F$102;"параметр значим ";"параметр незначим ")	
105	=Регрессия!D19	a2	=ЕСЛИ($B105>$F$102;"параметр значим ";"параметр незначим ")	
106				

Рисунок 3.24 – Проверка гипотез о статистической значимости оценок параметров модели на основе *F*- и *t*-критериев (формулы)

2 Проверка значимости оценок параметров модели по критерию Стьюдента.

Проверку гипотеза относительно каждого параметра проведем в соответствии со следующим алгоритмом.

Шаг 1. Формирование нулевой и альтернативной гипотез.

H_0: $\hat{a}_0 = 0$, $j = \overline{0;\ 2}$, – оценка *j*-го параметра является статистически незначимой, и j-й фактор никак не влияет на показатель *y*;

H_1: $\hat{a}_0 \neq 0,\quad j = \overline{0;\ 2}$, – оценка *j*-го параметра является статистически значащей, и *j*-й фактор никак влияет на показатель *y*.

Шаг 2. Выбор соответствующего уровня значимости.

Выбираем уровень значимости α = 0,05, т. е. доверительная вероятность P = 0,95.

Шаг 3. Выбираем расчетное значение *t*-критерия.

Для вычисления расчетных значений *t*-критерия в ячейках B103:B105 введем формулы (3.25) (см. рисунки 3.23 и 3.24). Для проверки полученных значений скопируем с итогового листа РЕГРЕССИЯ значения ячеек столбца *t*-статистика в ячейки C103:C105. Значения совпали.

Шаг 4. Определение по статистическим таблицам *t*-распределения Стьюдента критического значения *t*-критерия.

Критическое значение *t*-критерия находим по статистическим таблицам *t*-распределения Стьюдента по соответствующим значениям:

- доверительная вероятность P = 0,95;
- число степеней свободы k = 13.

Полученное табличное значение критерия $t_{\text{кр}}$ = 2,16 вводим в ячейку F102 или используем встроенную функцию СТЬЮДРАСПРОБ.

Шаг 5. Сравнение расчетного значения *t*-критерия с критическим и интерпретация результатов.

Выводы о принятии нулевой гипотезы, т. е. значимости оценок параметров модели сделаем в ячейках F103:F105, используя функцию ЕСЛИ. С надежностью Р = 0,95 можно считать, что:

- оценки 1-го и 2-го параметров модели значимы, т. е. оба фактора существенно влияют на показатель;
- оценка нулевого параметра не является статистически значимой.

VI Построение доверительных интервалов для оценок параметров модели.

Для построения доверительных интервалов для параметров модели выбираем уровень значимости α = 0,05, соответственно уровень доверия будет равен P = 0,95. Для каждого параметра вычисляем нижнюю и верхнюю границы доверительных интервалов по формулам (3.27) (рисунок 3.25).

	A	B	C	D	E
107		по формуле		регрессия	
108		нижние	верхние	нижние	верхние
109		-5,82	22,52	-5,82	22,52
110		0,12	0,22	0,12	0,22
111		4,72	11,63	4,72	11,63

а

	A	B	C	D	E
106					
107		по формуле		регрессия	
108		нижние	верхние	нижние	верхние
109		=W21-$F$102*B92	=W21+$F$102*B92	=Регрессия!F17	=Регрессия!G17
110		=W22-$F$102*B93	=W22+$F$102*B93	=Регрессия!F18	=Регрессия!G18
111		=W23-$F$102*B94	=W23+$F$102*B94	=Регрессия!F19	=Регрессия!G19

б

а – значения; б – формулы

Рисунок 3.25 – Доверительные интервалы для оценок параметров

Исходя из этого, 95%-е доверительные интервалы для параметров модели имеют вид

$$-5{,}82 \le a_0 \le 22{,}52,$$

$$-0{,}119 \le a_1 \le 0{,}216,$$

$$4{,}72 \le a_2 \le 11{,}63.$$

VII Прогнозирование на основе оцененной модели.

Поскольку оцененная модель является адекватной статистическим данным, то на основе этой модели можно осуществлять прогнозирование общих затрат для одного из предприятий объединения, деятельность которого исследуется.

1 Точечный прогноз индивидуального значения показателя.

Сделаем точечный прогноз общих затрат для одного из предприятий при условии, что грузооборот составит 200 тыс. усл. ден. ед., а фондоемкость – 3 тыс. усл. ден. ед., т. е. $x_{p1} = 200$, $x_{p2} = 3$, по формуле (3.29). Введем вектор прогнозных значений и, используя функцию МУМНОЖ, получим $\hat{y}_p = 66{,}38$ тыс. усл. ден. ед. (рисунки 3.26 и 3.27).

2 Доверительный интервал для прогноза математического ожидания общих затрат.

Рассчитываем значения нижней и верхней границ прогнозного интервала по формулам (3.31), т. е. $\hat{y}_p \pm t_{a/2} \cdot \hat{\sigma}_{\bar{y}_p}$.

Оценку дисперсии прогноза математического ожидания вычисляем как произведение трех матриц $\sigma^2_{\bar{y}_P} = X_P \cdot \hat{\Sigma}_{\hat{A}} \cdot X_P^T$ в следующей последовательности:

- используя функцию МУМНОЖ, находим произведение вектора прогнозных значений факторов и оцененной ковариационной матрицы оценок параметров модели (см. рисунки 3.26 и 3.27);
- при помощи функции ТРАНСП найдем транспонированный вектор прогнозных значений факторов в диапазоне H911:H121;
- используя функцию МУМНОЖ, найдем произведение матрицы $X_P \cdot \hat{\Sigma}_{\hat{A}}$ на матрицу X_P^T в ячейке F123.

	A	B	C	D	E	F	G	H
113					точечный прогноз общих затрат			
114			xp		8,347600109			
115		1	200	3	0,167529635	66,3793015		
116					8,175258118			
117		интервальный прогноз мат ожидания общих затрат						
118		ст ошибка прогноза мат ожидания						
119					43,03219304	-0,0061681	-6,589590711	1
120		1	200	3	-0,00616815	0,00049564	-0,024082769	200
121					-6,58959071	-0,0240828	2,556700129	3
122					1			
123		22,0298	0,020710601	-3,73604418	200	14,9637785		
124					3			
125		нижняя граница				**58,0237641**		
126		верхняя граница				**74,7348389**		
127					интервальный прогноз общих затрат			
128		ст ошибка прогноза				12,4318695		
129		нижняя граница				39,5264633		
130		верхняя граница				93,2321397		

Рисунок 3.26 – Прогнозирование по модели (значения)

	A	B	C	D
112				
113				точечный прогноз общи
114			xp	
115		1	200	3
116				
117		интервальный прогноз мат ожидания общих затрат		
118		ст ошибка прогноза мат ожидания		
119				
120		1	200	3
121				
122				
123		=МУМНОЖ(B120:D120;E119:G121)	=МУМНОЖ(B120:D120;E119:G121)	=МУМНОЖ(B120:D120;E119:G121)
124				
125		нижняя граница		
126		верхняя граница		
127				интервальный прогноз об
128		ст ошибка прогноза		
129		нижняя граница		
130		верхняя граница		

а

	E	F	G	H
112				
113	щих затрат			
114	=Регрессия!B17			
115	=Регрессия!B18	=МУМНОЖ(B115:D115;E114:E116)		
116	=Регрессия!B19			
117				
118				
119	=F86	=G86	=H86	=ТРАНСП(B120:D120)
120	=F87	=G87	=H87	=ТРАНСП(B120:D120)
121	=F88	=G88	=H88	=ТРАНСП(B120:D120)
122	=ТРАНСП(B120:D120)			
123	=ТРАНСП(B120:D120)	=МУМНОЖ(B123:D123;E122:E124)		
124	=ТРАНСП(B120:D120)			
125		**=F115-2,16*F123^0,5**		
126		**=F115+2,16*F123^0,5**		
127	общих затрат			
128		=(B83+F123)^0,5		
129		=F115-2,16*F128		
130		=F115+2,16*F128		

б

а – диапазон ячеек с A до D; б – диапазон ячеек с E до H

Рисунок 3.27 – Прогнозирование по модели (формулы)

Таким образом, 95%-й интервал доверия для прогноза математического ожидания общих затрат имеет вид $58{,}04737 \le \bar{y}_p \le 74{,}71138$, т. е. вероятность того, что $\bar{y}_p$ лежит в этом интервале, равна 0,95.

3 Доверительный интервал для прогноза индивидуального значения показателя.

Для нахождения интервального прогноза индивидуального значения общих затрат вычислим стандартную ошибку прогноза индивидуального значения по формуле (3.35). Таким образом, с вероятностью 0,95 можно утверждать, что прогнозное значение общих затрат принадлежит интервалу $39{,}53385 \le \bar{y}_p \le 93{,}2249$.

VIII Экономический анализ по оцененной модели.

Поскольку оцененная модель является адекватной статистическим данным, то на основе этой модели можно проводить экономический анализ исследуемого процесса. Для этого рассчитаем предельные и средние показатели (рисунки 3.28 и 3.29).

	A	B	C	D	E	F	G	H
131								
132				средняя эффективность показателя	граничная эффективность показателя	частичная эластичность общих затрат	суммарная эластичность	граничная норма замещения показателя
133		грузооборот x1		0,360413589	0,167529635	0,464826078	0,908769398	-51,09434
134		фондоемкость x2		18,41509434	8,175258118	0,44394332		-0,0195716

Рисунок 3.28 – Расчет средних и предельных показателей (значения)

	A	B	D	E	F	G	H
132			средняя эффективность показателя	граничная эффективность показателя	частичная эластичность общих затрат	суммарная эластичность	граничная норма замещения показателя
133		грузооборот x1	=C20/D20	=W22	=E133/D133		=-D134/D133
134		фондоемкость x2	=C20/E20	=W23	=E134/D134	=F133+F134	=-D133/D134
135							
136							

Рисунок 3.29 – Расчет средних и предельных показателей (формулы)

Анализ полученных результатов приводит к следующему выводу:

- на основании значения средней эффективности грузооборота можно утверждать, что на 1 тыс. усл. ден. ед. грузооборота в среднем приходится 0,36 тыс. усл. ден. ед. общих затрат;

- на основании значения средней эффективности фондоемкости можно утверждать, что на 1 тыс. усл. ден. ед. фондоемкости в среднем приходится 18,42 тыс. усл. ден. ед. общих затрат;

- на основании значения предельной эффективности грузооборота можно утверждать, что при увеличении грузооборота на 1 тыс. усл. ден. ед. объем общих затрат увеличится на 0,17 тыс. усл. ден. ед. при неизменном объеме фондоемкости;

- на основании значения предельной эффективности фондоемкости можно утверждать, что при увеличении фондоемкости на 1 тыс. усл. ден. ед. объем общих затрат увеличится на 8,18 тыс. усл. ден. ед. при неизменном объеме грузооборота;

- на основании значения коэффициента частичной эластичности по фактору X_1 можно утверждать, что при увеличении грузооборота на 1 % объем общих затрат увеличится на 0,47 % при неизменном объеме фондоемкости;

- на основании значения коэффициента частичной эластичности по фактору X_2 можно утверждать, что при увеличении фондоемкости на 1 % объем общих затрат увеличится на 0,44 % при неизменном объеме грузооборота;

- на основании значения предельной нормы замещения 2-го фактора 1-м можно утверждать, что для замещения 1 тыс. усл. ден. ед. фондоемкости будет необходимо 51,09 тыс. усл. ден. ед. грузооборота при сохранении неизменного уровня общих затрат;

- на основании значения предельной нормы замещения 1-го фактора 2-м можно утверждать, что для замещения 1 тыс. усл. ден. ед. грузооборота будет необходимо 0,02 тыс. усл. ден. ед. фондоемкости при сохранении неизменного уровня общих затрат.

3.13 Варианты заданий

По статистическим данным девяти предприятий общественного питания за год построить линейную двухфакторную эконометрическую модель, которая характеризует зависимость между уровнем рентабельности (%), относительным уровнем издержек обращения (%) и трудоемкостью (численностью работников в расчете на 1 тыс. усл. ден. ед. товарооборота). Оценить адекватность модели. Выполнить прогноз уровня рентабельности для одного из предприятий (прогнозные значения факторов выбрать самостоятельно). Сделать экономический анализ характеристик взаимосвязей. Выходные данные приведены в таблицах-вариантах 1–24 (данные условные).

Примечание – Задачи выполнены в соответствии с алгоритмом, приведенным в методических указаниях.

Вариант 1

№	Рентабель-ность	Издержки оборота	Трудоем-кость
1	2,32	38,8	114
2	2,19	39,9	103,1
3	2,83	30,1	103,8
4	2,75	31,7	146
5	2,59	37,2	124,8
6	2,27	39,7	103,6
7	2,05	36,9	119
8	1,95	38,2	108,7
9	2,08	40,1	106,5

Вариант 2

№	Рентабель-ность	Издержки оборота	Трудоем-кость
1	2,58	15,1	120
2	2,64	16,1	118,4
3	2,52	16,7	108,4
4	2,75	15,4	110
5	2,63	17,1	105,9
6	2,48	16,8	117,7
7	2,62	16,9	97,5
8	2,88	16,1	113,7
9	2,68	15	122,3

Вариант 3

№	Рентабель-ность	Издержки оборота	Трудоем-кость
1	2,48	16,8	117,7
2	2,62	16,9	97,5
3	2,88	16,1	113,7
4	2,68	15	122,3
5	2,52	18	102
6	2,74	17,2	106,7
7	2,56	17,1	108,5
8	2,68	16,4	114,3
9	2,55	16,7	94,3

Вариант 4

№	Рентабель-ность	Издержки оборота	Трудо-емкость
1	2,75	16,8	110
2	2,63	15,5	105,9
3	2,48	17	117,7
4	2,62	16,8	97,5
5	2,88	16,9	113,7
6	2,68	16,1	122,3
7	2,56	15	102
8	2,74	18	106,7
9	2,6	17,2	108,5

Вариант 5

№	Рентабель-ность	Издержки оборота	Трудоем-кость
1	2,92	14,1	87,8
2	2,64	17,2	72
3	2,79	17,1	72,4
4	2,67	17,8	69,5
5	2,68	16,2	75
6	2,85	17,2	70,6
7	2,4	16,8	73,4
8	2,91	14,8	80,7
9	2,29	19,6	62,2

Вариант 6

№	Рентабель-ность	Издержки оборота	Трудо-емкость
1	2,45	17,1	71,3
2	2,38	19,5	61,7
3	3,04	12,5	96,2
4	2,67	16,5	72,9
5	2,7	16	75
6	2,65	16,1	74,6
7	2,79	16,2	74,1
8	2,49	18	66,9
9	3,27	11,4	98,6

Вариант 7

№	Рентабельность	Издержки оборота	Трудоемкость
1	2,85	17,2	70,6
2	2,4	16,8	73,4
3	2,91	14,8	80,7
4	2,29	19,6	62,2
5	3,27	11,4	98,6
6	2,45	17,1	71,3
7	2,38	19,5	61,7
8	3,04	12,5	96,2
9	2,67	16,5	72,9

Вариант 8

№	Рентабельность	Издержки оборота	Трудоемкость
1	2,93	18,1	71,2
2	2,5	17,2	73,4
3	2,95	14,9	81,2
4	2,39	20,1	63,7
5	3,25	11,4	96,6
6	2,65	17,1	72,2
7	2,42	19,5	61,7
8	3,14	17,5	96,2
9	2,75	16,7	72,9

Вариант 9

№	Рентабельность	Издержки оборота	Трудоемкость
1	2,14	17,4	80,3
2	2,94	13,8	102,5
3	2,67	15	94,3
4	2,44	18,6	76
5	2,83	16,2	87,3
6	2,92	15,7	96,1
7	2,61	17,9	82,8
8	2,72	15,3	96,9
9	2,68	16,3	83,7

Вариант 10

№	Рентабельность	Издержки оборота	Трудоемкость
1	2,67	15	94
2	2,45	18,6	78
3	2,86	16,2	87,5
4	2,9	15,7	90,2
5	2,6	17,9	84,8
6	2,72	16,3	95,9
7	2,68	17,7	91
8	2,5	16,8	84,7
9	2,74	17,5	88,2

Вариант 11

№	Рентабельность	Издержки оборота	Трудоемкость
1	2,83	13,8	68
2	2,75	14,8	64,3
3	2,4	16,9	55,1
4	2,3	16,8	55,5
5	2,47	14,8	63,3
6	2,45	17,9	52,7
7	2,48	17,6	53,7
8	2,41	15,7	60,2
9	2,34	15,2	62,2

Вариант 12

№	Рентабельность	Издержки оборота	Трудоемкость
1	2,69	14,9	69,4
2	2,48	16,1	58,7
3	2,11	19,7	62,3
4	2,82	14	83,8
5	2,43	17,1	68,5
6	2,34	18,2	64,5
7	2,48	17,4	67,6
8	2,69	16,1	72,9
9	2,36	18,8	62,4

Вариант 13

№	Рентабель-ность	Издержки оборота	Трудоем-кость
1	2,35	16,8	56,5
2	2,48	15	64,3
3	2,45	17,5	53,7
4	2,47	17,6	54,7
5	2,42	15,7	60,2
6	2,34	15,2	62,4
7	2,7	14,9	69,5
8	2,48	16,1	58,7
9	2,15	19,7	62,3

Вариант 14

№	Рентабель-ность	Издержки оборота	Трудо-емкость
1	2,46	19	65,4
2	2,7	16,3	73,9
3	2,58	17,5	68,5
4	2,34	18,4	64,5
5	2,43	17,2	69,3
6	2,84	15	83,8
7	2,15	19,8	62,3
8	2,49	16,1	58,7
9	2,6	14,9	69,4

Вариант 15

№	Рентабель-ность	Издержки оборота	Трудоем-кость
1	2,27	32,1	112,5
2	1,94	31	116,4
3	2,32	32,4	111,6
4	2,49	33,2	108,9
5	2,57	31,2	116,5
6	2,01	34,8	104,5
7	1,87	35,4	102,7
8	2,39	33	110,2
9	2,18	34,8	104,7

Вариант 16

№	Рентабель-ность	Издержки оборота	Трудо-емкость
1	2,17	33,3	109,4
2	1,8	36,1	101,1
3	2,36	38,3	102,6
4	2,5	30,6	128,5
5	2,57	32,1	122,5
6	2,33	37,6	105,2
7	2,51	34,8	114,8
8	2,4	34,2	116
9	2,5	34,2	116

Вариант 17

№	Рентабель-ность	Издержки оборота	Трудоем-кость
1	2,08	32,5	81,6
2	1,99	33,4	79,4
3	1,96	37,8	69,5
4	2,18	35,8	85,4
5	1,91	34,2	84,3
6	2,37	37,2	71,4
7	1,92	38,2	78,1
8	2,15	29,4	90,8
9	2,41	37,2	72,1

Вариант 18

№	Рентабель-ность	Издержки оборота	Трудо-емкость
1	2,32	38	114,5
2	2,19	39	103
3	2,83	30	120
4	2,75	31	145
5	2,59	37	124
6	2,27	39	103
7	2,05	36	119,5
8	1,95	38	104,5
9	2,08	40	106,5

Вариант 19

№	Рентабель-ность	Издержки оборота	Трудоем-кость
1	2,95	29,4	122,5
2	2,55	35,4	126,2
3	2,26	29,7	112,6
4	2,49	37,1	120,5
5	2,17	35,7	125,3
6	2,38	40,2	111,3
7	2,22	39,4	112,2
8	2,64	43,7	121,2
9	2,63	38,4	126,4

Вариант 20

№	Рентабель-ность	Издержки оборота	Трудо-емкость
1	2,2	34,5	83,6
2	1,96	35,0	76,5
3	2,25	43,7	76,9
4	2,69	31,9	104,6
5	2,24	37,3	72,3
6	2,43	40,9	66,3
7	2,32	38,8	69,6
8	2,6	35,7	75,6
9	2,7	43,2	62,4

Вариант 21

№	Рентабель-ность	Издержки оборота	Трудоем-кость
1	2,93	3,10	10,24
2	5,27	9,8	7,51
3	6,85	12,1	10,81
4	7,01	12,9	9,89
5	7,02	11,2	13,72
6	8,35	14,9	13,92
7	4,33	7,8	8,54
8	5,77	9,4	12,36
9	7,68	12,9	12,27

Вариант 22

№	Рентабель-ность	Издержки оборота	Трудо-емкость
1	1,30	6,6	42,2
2	1,37	6,7	43,8
3	1,47	6,7	45,6
4	1,59	6,8	47,6
5	1,69	6,9	50,3
6	1,82	7,0	53,1
7	1,95	7,1	55,9
8	1,98	7,0	58,8
9	1,93	7,0	61,6

Вариант 23

№	Рентабель-ность	Издержки оборота	Трудоем-кость
1	2,8	11,5	25
2	3,0	14,5	26
3	3,7	16,6	26
4	3,5	16,4	24
5	3,6	17,9	25
6	3,8	15,5	28
7	4,0	18,4	30
8	4,5	20,3	32
9	4,6	21,8	35

Вариант 24

№	Рентабель-ность	Издержки оборота	Трудо-емкость
1	2,17	33,3	10,94
2	1,80	36,1	10,11
3	2,36	38,3	10,26
4	2,50	30,6	12,85
5	2,57	32,1	12,25
6	2,33	37,6	10,52
7	2,51	34,8	11,48
8	2,40	34,2	11,60
9	2,50	34,2	11,60

Практическая работа №4
Исследование остатков эконометрической модели на гетероскедастичность

4.1 Понятие обобщенной модели. Обобщенный метод наименьших квадратов

Классическая модель в определенных случаях требует обобщения, чтобы быть пригодной для эмпирических экономических и социальных исследований. Наиболее распространенными случаями нарушения классической модели являются:

- гетероскедастичность остатков;
- автокорреляция остатков во временных рядах;
- пространственная корреляция остатков в пространственных рядах.

Если предпосылки классической модели нарушены, то 1МНК-оценщики часто не имеют желаемых статистических свойств, а именно, оценка вектора $\hat{A}$ является несмещенной, обоснованной, но неэффективной. В этом случае классическую модель регрессии обобщают, а методы оценки модифицируют.

Под ***обобщенной линейной регрессионной моделью*** понимают, как правило, такую классическую регрессионную модель, которая обобщена только по одному конкретному условию – по ковариационной матрице вектора остатков, которая в классической модели является диагональной, с равнозначными элементами на диагонали, а в обобщенной модели она может не являться диагональной и иметь на диагонали элементы разной величины. Общая форма ковариационной матрицы вектора остатков имеет вид

$$\Sigma_e = \begin{pmatrix} \sigma_{e_1}^2 & \sigma_{e_1 e_2} & \dots & \sigma_{e_1 e_n} \\ \sigma_{e_2 e_1} & \sigma_{e_2}^2 & \dots & \sigma_{e_2 e_n} \\ \dots & \dots & \dots & \dots \\ \sigma_{e_n e_1} & \sigma_{e_n e_2} & \dots & \sigma_{e_n}^2 \end{pmatrix}, \text{ т. е. } \Sigma_e = \sigma_e^2 \cdot S, \qquad (4.1)$$

где $\sigma_{e_i}^2$ – дисперсия остатков;

S – матрица общего вида.

Форма ковариационной матрицы вектора остатков в классической модели имеет вид

$$\Sigma_e = \begin{pmatrix} \sigma_e^2 & 0 & \dots & 0 \\ 0 & \sigma_e^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_e^2 \end{pmatrix}, \text{ т. е. } \Sigma_e = \sigma_e^2 \cdot E, \tag{4.2}$$

где E – единичная матрица, т. е. в классической модели, $S = E$, что нереально в практических исследованиях.

Обобщенная оценка методом наименьших квадратов (***ОМНК***), которая называется также ***оценкой Эйткена***, приводит при известной ковариационной матрице $\Sigma_e = \sigma_e^2 \cdot S$ к следующим модифицированным формулам оценки (таблица 4.1).

Таблица 4.1 – Формулы оценивания 1МНК и ОМНК

Название	**Одношаговый метод наименьших квадратов (1МНК)**	**Обобщенный метод наименьших квадратов (ОМНК)**
Оператор оценивания	$\hat{A} = \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y$	$\hat{A} = \left(X^T \cdot S^{-1} \cdot X\right)^{-1} \cdot X^T \cdot S^{-1} \cdot Y$
Оценка дисперсии остатков	$\hat{\sigma}_e^2 = \dfrac{e^T \cdot e}{n-m-1}$	$\hat{\sigma}_e^2 = \dfrac{e^T \cdot S^{-1} \cdot e}{n-m-1}$
Оценка ковариационной матрицы оценок параметров	$\hat{\Sigma}_{\hat{A}} = \hat{\sigma}_e^2 \cdot \left(X^T \cdot X\right)^{-1}$	$\hat{\Sigma}_{\hat{A}} = \hat{\sigma}_e^2 \cdot \left(X^T \cdot S^{-1} \cdot X\right)^{-1}$
Коэффициент множественной детерминации	$R^2 = 1 - \dfrac{e^T \cdot e}{Y^T \cdot Y - n \cdot \bar{y}^2}$	$R^2 = 1 - \dfrac{e^T \cdot S^{-1} \cdot e}{Y^T \cdot Y - n \cdot \bar{y}^2}$

Оценки Эйткена в обобщенной модели не имеют недостатков, присущих оценкам, полученным на основе 1МНК для этой модели, но при обязательном условии, что *ковариационная матрица известна*.

4.2 Гетероскедастичность

4.2.1 Понятие гомоскедастичности и гетероскедастичности. Рассмотрим специфику эконометрического моделирования при возникновении предположения о гомоскедастичности остатков.

Если дисперсия остатков постоянна для каждого наблюдения, т. е. ковариационная матрица остатков имеет вид (4.2): $\Sigma_e = \sigma_e^2 \cdot E$, то это явление называется гомоскедастичность. Классическая модель с постоянными дисперсиями называется гомоскедастичной.

Если дисперсия остатков изменяется для каждого наблюдения или группы наблюдений, т. е. ковариационная матрица остатков имеет вид (4.1): $\Sigma_e = \sigma_e^2 \cdot S$, где S – диагональная матрица, то это явление называется гетероскедастичностью.

Например, при построении эконометрической модели, характеризующей зависимость между сбережениями и доходами населения, можно выдвинуть гипотезу, что дисперсия остатков по отдельным группам населения будет изменяться и будет пропорциональна среднему доходу этой группы.

Поскольку явление гетероскедастичности связано только с тем, что меняются дисперсии остатков, то матрица должна быть положительно определенной и диагональной, но не все ее диагональные элементы равнозначны:

$$S = \begin{pmatrix} \frac{1}{\lambda_1} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{\lambda_2} & 0 & \dots & 0 \\ 0 & 0 & \frac{1}{\lambda_3} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \frac{1}{\lambda_n} \end{pmatrix},$$

т. е. обратная матрица $$S^{-1} = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}. \qquad (4.3)$$

В этой матрице значения λ_i можно вычислить тремя способами, в зависимости от того, какая гипотеза выдвинута относительно изменения дисперсии остатков:

- если дисперсия остатков пропорциональна изменению фактора, т. е. $\Sigma_e = \sigma_e^2 \cdot x_{ij}$, то $\lambda_i = \frac{1}{x_{ij}}$ $\left(i = \overline{1;n},\ j = \overline{1;m}\right)$;

- если дисперсия остатков пропорциональна изменению квадрата фактора, т. е. $\Sigma_e = \sigma_e^2 \cdot x_{ij}^2$, то $\lambda_i = \frac{1}{x_{ij}^2}$ $\left(i = \overline{1;n},\ j = \overline{1;m}\right)$;

- если дисперсия остатков пропорциональна изменению квадратов модулей остатков, т. е. $\Sigma_e = \sigma_e^2 \cdot \{|e_i|\}^2$, то $\lambda_i = \{|e_i|\}^2 \ \left(i = \overline{1;n}\right)$.

После определения оценок параметров модели операторами оценки ОМНК необходимо провести исследование остатков. Если теперь их колебания имеют неупорядоченный характер, то это может свидетельствовать о том, что коррекция на гетероскедастичность прошла успешно. Конечно, необходимо сравнить и вторые параметры регрессии (значимость оценок, сумма квадратов остатков и др.) И только тогда принимать решение о том, какая из моделей, оцененая 1МНК или ОМНК, наиболее пригодна.

4.2.2 Методы исследования гетероскедастичности. Опишем несколько наиболее распространенных статистических тестов на гетероскедастичность без проведения их детального исследования (исключая тесты Гольдфельда – Квандта). Во всех этих тестах проверяется нулевая гипотеза H_0: $\sigma_{e_1}^2 = \sigma_{e_2}^2 = \ldots = \sigma_{e_n}^2 = \sigma_e^2 = \text{const}$, против альтернативной гипотезы H_A: $\sigma_{e_i}^2 \neq \text{const}\left(i = \overline{1;n}\right)$.

1 Тест Уайта (критерий μ).

Используется в случае достаточно большой совокупности наблюдений (подробно рассмотрено в работе [5]).

2 Параметрический тест Гольдфельда – Квандта.

Используется, как правило, в тех случаях, когда априорно предполагается, что дисперсии остатков зависят от величины одного из факторов модели и совокупность наблюдений невелика.

Алгоритм тестирования Гольдфельда – Квандта

Шаг 1. Формулирование нулевой и альтернативной гипотез.

H_0: $\sigma_{e_1}^2 = \sigma_{e_2}^2 = \ldots = \sigma_{e_n}^2 = \sigma_e^2 = \text{const}$, т. е. дисперсия остатков постоянна, остатки являются гомоскедастичными.

H_A: $\sigma_{e_i}^2 \neq \text{const}\left(i = \overline{1;n}\right)$, т. е. дисперсия остатков изменяется, остатки являются гетероскедастичными.

Шаг 2. Упорядочение наблюдений по росту элементов вектора X_j.

Наблюдения упорядочиваются в соответствии с величиной элементов вектора X_j – фактора, который может вызвать гетероскедастичность.

Шаг 3. Отбрасывания наблюдений с, которые содержатся в середине вектора наблюдений.

На основе экспериментальных расчетов авторы теста получили оптимальное соотношение между количеством наблюдений c и объемом выборочной совокупности n:

$$\frac{c}{n} = \frac{4}{15}, \qquad (4.4)$$

т. е. количество наблюдений, которые нужно отбросить, рассчитывается по формуле

$$c = \frac{4}{15} \cdot n. \qquad (4.5)$$

Шаг 4. Построение эконометрических моделей по методу 1МНК для двух полученных совокупностей наблюдений.

Построить по методу 1МНК две эконометрические модели по двум полученным совокупностями наблюдений объемом:

- $n_1 = n_2 = \dfrac{n-c}{2}$ в случае парных значений наблюдений c и n;

- $n_1 \neq n_2$ в случае нечетных значений наблюдений c или n – тест обобщается при условии, что $n_1, n_2 > m+1$.

Шаг 5. Вычисление статистики F.

Значение статистики F исчисляется:

- как отношение сумм квадратов отклонений второй и первой моделей:

$$F = \frac{SSE_2}{SSE_1}, \text{ если } n_1 = n_2; \qquad (4.6)$$

- как отношение средних взвешенных сумм квадратов отклонений второй и первой моделей:

$$F = \frac{MSE_2}{MSE_1}, \text{ если } n_1 \neq n_2. \qquad (4.7)$$

Если гипотеза H_0 верна, то F имеет распределение Фишера со $[n_1 - (m+1); n_2 - (m+1)]$ степенями свободы.

Шаг 6. Определение по статистическим таблицам F-распределения Фишера критического значения F-критерия.

Критическое значение F-критерия находят по статистическим таблицам F-распределения Фишера по соответствующим значениям:

- доверительной вероятности P;

- степеней свободы $k_1 = n_1 - m - 1$ и $k_2 = n_2 - m - 1$.

Если так, то нет оснований отклонять нулевую гипотезу о том, что дисперсия остатков постоянна для каждого наблюдения, т. е. с принятой надежностью можно утверждать, что гетероскедастичность остатков отсутствует – остатки гомоскедастичны.

Если нет, то нулевая гипотеза отклоняется, т. е. с принятой надежностью можно утверждать, что существует гетероскедастичность остатков.

Шаг 7. Сравнение расчетного значения отношения F с критическим и интерпретация результатов.

Если $F < F_{\text{кр}}$, то нет оснований отклонять нулевую гипотезу о том, что дисперсия остатков однородна для каждого наблюдения, т. е. с принятой надежностью можно утверждать, что гетероскедастичность остатков отсутствует – остатки гомоскедастичны.

Если $F > F_{\text{кр}}$, то нулевая гипотеза отклоняется, т. е. с принятой надежностью можно утверждать, что существует гетероскедастичность остатков.

3 Непараметрический тест Гольдфельда – Квандта.

Этот тест основан на анализе графиков остатков. Если существует предположение о зависимости остатков от одного из факторов, то целесообразно упорядочить наблюдения в порядке возрастания значений этого фактора, построить регрессию и графики остатков:

- если для всех значений фактора $X_j\left(j=\overline{1;m}\right)$ остатки распределяются примерно одинаково (это видно при обычном имеющемся исследовании), то дисперсия их однородна, т. е. имеет место гомоскедастичность;

- если разброс колебаний также возрастает, то имеет место гетероскедастичность.

Этот тест достаточно простой, но не такой надежный, как параметрический тест.

4 Тест Глейсер.

Этот тест основан на построении регрессии, характеризующей зависимость величины остатков по модулю от фактора X_j, который может вызвать гетероскедастичность.

5 Тест Бреуша – Пагана.

Этот тест используется в тех случаях, когда априорно предполагается, что дисперсии остатков зависят от некоторых дополнительных факторов.

4.2.3 Прогнозирование по обобщенной модели. Когда параметры модели оцениваются по ОМНК, проблема прогнозирования требует специального исследования. Пусть имеем модель

$$Y = X \cdot \hat{A} + u\,,$$

где $u = Y - X \cdot \hat{A}$ – отклонения по обобщенной модели, такие, что

$$M(u) = 0 \text{ и } M\left(u \cdot u^T\right) = \sum{}_u = \sigma_u^2 \cdot S = V.$$

Задача сводится к тому, чтобы предсказать значение зависимой переменной Y_p для заданного вектора X_p. Можно записать

$$Y_p = X_p \cdot A + u_p, \tag{4.8}$$

где u_p – неизвестное значение отклонений в прогнозный период.

Пусть для u_p имеем

$$M(u_p) = 0 \text{ и } M\left(u_p \cdot u_p{}^T\right) = \sum{}_{uP} = \sigma_u^2 \cdot E, \tag{4.9}$$

и вектор ковариаций текущих и прогнозных значений отклонений

$$W = M\left(u_p, u\right) = \begin{pmatrix} M\left(u_1, u_p\right) \\ M\left(u_2, u_p\right) \\ \dots \\ M\left(u_n, u_p\right) \end{pmatrix}. \tag{4.10}$$

Для прогноза можно использовать соотношение

$$\hat{Y}_p = X_p \cdot \hat{A} + W^T \cdot V^{-1} \cdot u_p, \tag{4.11}$$

где $X_p \cdot \hat{A}$ – оценка прогнозной величины;

$W^T \cdot V^{-1} \cdot u_p$ – отклонение прогноза.

Этот прогноз имеет две особенности:

- вектор прогнозных значений X_p умножается на вектор прогнозных оценок, вычисленных согласно ОМНК;

- для оценки неизвестных прогнозных остатков u_p применяется матрица V, содержащая информацию о взаимозависимости остатков базисного периода.

4.3 Алгоритм исследования наличия гетероскедастичности и оценки параметров модели обобщенного метода Эйткена в среде табличного процессора MS Excel

I Спецификация модели.

1 Идентификация переменных.

2 Общий вид линейной модели и ее оценки.

II Оценка параметров модели по методу 1МНК.

III Исследование наличия гетероскедастичности в массиве исходных данных по непараметрическому тесту Гольдфельда – Квандта.

IV Исследование наличия гетероскедастичности в массиве исходных данных по параметрическому тесту Гольдфельда – Квандта.

V Оценка параметров модели ОМНК (методом Эйткена).

1 Определение матриц, входящих в оператор Эйткена.

2 Вычисление произведений матриц.

3 Определение обратной матрицы $\left(X^T S^{-1} X\right)^{-1}$ и вектора $X^T S^{-1} Y$.

4 Вычисление вектора оценок параметров модели.

VI Сравнительный анализ оценок параметров, полученных разными методами, выводы о смещенности оценок параметров моделей.

VII Проверка существенности связи, описываемой моделями, параметры которых оценены по методу Эйткена и методом 1МНК.

4.4 Исследование остатков эконометрической модели на гетероскедастичность в оболочке электронных таблиц MS Excel

По статистическим данным о годовой сумме сбережений и годовом доходе жителей (семей) города необходимо проверить гипотезу о наличии гетероскедастичности, построить эконометрические модели по методу Эйткена и одношаговым методом наименьших квадратов, сравнить модели и сделать выводы о точности оценок. Выборка из 18-ти разных семей $(n=18)$ приведена в таблице 4.1.

Поставленную задачу решим в среде Excel по следующему алгоритму.

I Спецификация модели.

1 Идентификация переменных.

Y – годовая сумма сбережений жителей города – результирующий показатель;

X – годовой доход жителей города – показатель-фактор.

Блок данных формируется в массиве **C3:C20**, в ячейки **B3:B20** вводятся номера наблюдений (рисунок 4.1, а).

2 Общий вид линейной однофакторной модели и ее оценки.

Предположим, что между сбережениями жителей города и их доходом существует зависимость, которая может быть описана линейной однофакторной моделью $y = a_0 + a_1 x + e$, ее оценкой является уравнение $\hat{y} = \hat{a}_0 + \hat{a}_1 x$.

II Оценка параметров модели по методу 1МНК.

С целью исследования массива данных на гетероскедастичность остатков по непараметрическому тесту Гольдфельда – Квандта построим модель по методу 1МНК (исходные данные упорядочены). Алгоритм построения линейной эконометрической модели в оболочке электронных таблиц Excel приведен в [7]. Рассмотрим отдельно, как задавать параметры окна диалога **Регрессия**.

	B	C	D	E	F	G	H
2	№ п/п	Сбережения	Доход	Предсказанное сбережение	Остатки	Оцененное Y, ОМНК	Остатки, e
3	1	2,3	15	2,2145	0,0855	2,2295	0,0705
4	2	2,2	15	2,2145	-0,0145	2,2295	-0,0295
5	3	2,08	16	2,2288	-0,1488	2,2434	-0,1634
6	4	2,2	17	2,2431	-0,0431	2,2572	-0,0572
7	5	2,1	17	2,2431	-0,1431	2,2572	-0,1572
8	6	2,32	18	2,2574	0,0626	2,2710	0,0490
9	7	2,45	19	2,2717	0,1783	2,2848	0,1652
10	8	2,5	20	2,2860	0,2140	2,2986	0,2014
11	9	2,2	20	2,2860	-0,0860	2,2986	-0,0986
12	10	2,5	22	2,3146	0,1854	2,3262	0,1738
13	11	3,1	64	2,9150	0,1850	2,9065	0,1935
14	12	2,4	68	2,9722	-0,5722	2,9618	-0,5618
15	13	2,82	72	3,0294	-0,2094	3,0170	-0,1970
16	14	3,04	80	3,1438	-0,1038	3,1276	-0,0876
17	15	2,7	85	3,2153	-0,5153	3,1966	-0,4966
18	16	3,91	90	3,2868	0,6232	3,2657	0,6443
19	17	3,1	95	3,3583	-0,2583	3,3348	-0,2348
20	18	3,99	100	3,4297	0,5603	3,4039	0,5861
21	Расчетное значение С				4,8	Принимаем	C=4

а

	B	C	D	E	F	G	H
2	№ п/п	Сбережения	Доход	Предсказанное сбережение	Остатки	Оцененное Y, ОМНК	Остатки, e
3	1	2,3	15	='1МНК'!B24	='1МНК'!C24	=$K$68+D3*$K$69	=C3-G3
4	2	2,2	15	='1МНК'!B25	='1МНК'!C25	=$K$68+D4*$K$69	=C4-G4
5	3	2,08	16	='1МНК'!B26	='1МНК'!C26	=$K$68+D5*$K$69	=C5-G5
6	4	2,2	17	='1МНК'!B27	='1МНК'!C27	=$K$68+D6*$K$69	=C6-G6
7	5	2,1	17	='1МНК'!B28	='1МНК'!C28	=$K$68+D7*$K$69	=C7-G7
8	6	2,32	18	='1МНК'!B29	='1МНК'!C29	=$K$68+D8*$K$69	=C8-G8
9	7	2,45	19	='1МНК'!B30	='1МНК'!C30	=$K$68+D9*$K$69	=C9-G9
10	8	2,5	20	='1МНК'!B31	='1МНК'!C31	=$K$68+D10*$K$69	=C10-G10
11	9	2,2	20	='1МНК'!B32	='1МНК'!C32	=$K$68+D11*$K$69	=C11-G11
12	10	2,5	22	='1МНК'!B33	='1МНК'!C33	=$K$68+D12*$K$69	=C12-G12
13	11	3,1	64	='1МНК'!B34	='1МНК'!C34	=$K$68+D13*$K$69	=C13-G13
14	12	2,4	68	='1МНК'!B35	='1МНК'!C35	=$K$68+D14*$K$69	=C14-G14
15	13	2,82	72	='1МНК'!B36	='1МНК'!C36	=$K$68+D15*$K$69	=C15-G15
16	14	3,04	80	='1МНК'!B37	='1МНК'!C37	=$K$68+D16*$K$69	=C16-G16
17	15	2,7	85	='1МНК'!B38	='1МНК'!C38	=$K$68+D17*$K$69	=C17-G17
18	16	3,91	90	='1МНК'!B39	='1МНК'!C39	=$K$68+D18*$K$69	=C18-G18
19	17	3,1	95	='1МНК'!B40	='1МНК'!C40	=$K$68+D19*$K$69	=C19-G19
20	18	3,99	100	='1МНК'!B41	='1МНК'!C41	=$K$68+D20*$K$69	=C20-G20
21	Расчетное значение С				4,8	Принимаем	C=4

б

а – значения; б – формулы

Рисунок 4.1 – Исходные данные. Оценки 1МНК и ОМНК

Входной интервал Y: вводим адрес диапазона зависимых переменных **C3:C20**.

Входной интервал X : введем адрес диапазона независимых переменных **D3:D20**.

Параметры вывода: установим флажок на **Новый рабочий лист**, чтобы открыть новый лист в книге и вставить результаты анализа, начиная с ячейки **A1**, введем имя нового листа «1МНК» в поле, расположенное напротив.

Остатки: установить флажок, чтобы включить остатки в итоговый диапазон.

График остатков: установить флажок, чтобы построить диаграмму остатков для каждой независимой переменной.

График подбора: установить флажок, чтобы построить диаграммы наблюдаемых и оценочных значений.

Остальные параметры окна диалога **Регрессия** нет необходимости активизировать. Нажимаем клавишу **Enter** или кнопку **OK**.

Скопируем с листа «1МНК» значение таблицы **Вывод остатка** в ячейки **E2:F2O** (см. рисунок 4.1); значения ячеек столбца *Коэффициенты* строк **Y-пересечение** и **Доход** – в ячейки B25:C27 (рисунок 4.2).

	B	C
24	1МНК	
25		*Коэффициенты*
26	Y-пересечение	2,000019171
27	Доход	0,014297305
28	Модель 1	
29		*Коэффициенты*
30	*Y-пересечение*	*1,474680851*
31	*Переменная X 1*	*0,045531915*
32	Модель 2	
33		*Коэффициенты*
34	*Y-пересечение*	*-0,259545298*
35	*Переменная X 1*	*0,04029969*

а

	B	C
24	1МНК	
25		*='1МНК'!B16*
26	='1МНК'!A17	='1МНК'!B17
27	='1МНК'!A18	='1МНК'!B18
28	Модель 1	
29		*=Модель1!B16*
30	*=Модель1!A17*	*=Модель1!B17*
31	*=Модель1!A18*	*=Модель1!B18*
32	Модель 2	
33		*=Модель2!B16*
34	*=Модель2!A17*	*=Модель2!B17*
35	*=Модель2!A18*	*=Модель2!B18*

б

а – значения; б – формулы

Рисунок 4.2 – Построение моделей: инструмент **Регрессия**

Таким образом, эконометрическая модель сбережений жителей города, которая оценена по 1МНК, имеет вид

$$\hat{y} = 2{,}000019 + 0{,}014297x.$$

III Исследование наличия гетероскедастичности в массиве исходных данных по непараметрическому тесту Гольдфельда – Квандта.

Очень вероятно, что в рассматриваемой модели остатки будут гетероскедастичными, потому что сбережения семей зависят от объемов годовых доходов и величин расходов, что обуславливает разные значения дисперсий остатков.

На основании имеющегося анализа графиков подбора и остатков можно сделать вывод о наличии гетероскедастичности, поскольку дисперсии остатков для различных групп семей значительно отличаются (рисунки 4.3 и 4.4).

Можно принять, что:

- $\sigma_e^2 = \sigma_1^2$ – для семей с доходом $x_i < 50$ (группа 1);

- $\sigma_e^2 = \sigma_2^2$ – для семей с доходом $x_i > 50$ (группа 2).

Примечание – При анализе многофакторной модели на основании имеющегося исследования графиков подбора и остатков необходимо сделать предварительный вывод о наличии гетероскедастичности остатков и определить фактор, который может ее вызвать.

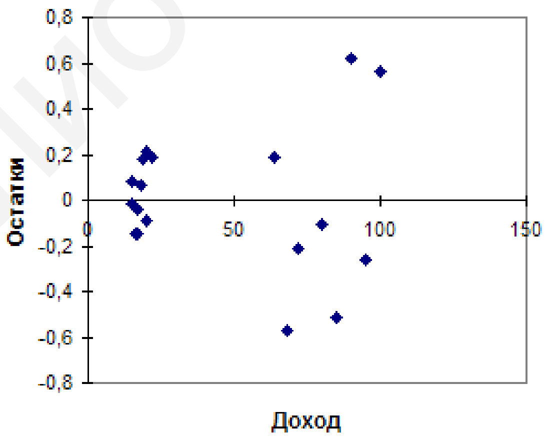


Рисунок 4.3 – График остатков (Доход)

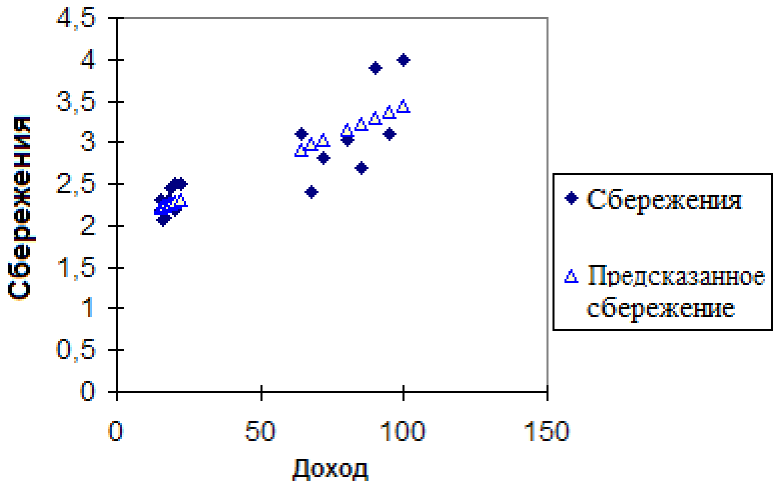


Рисунок 4.4 – График подбора

IV Исследование наличия гетероскедастичности в массиве исходных данных по параметрическому тесту Гольдфельда – Квандта.

Шаг 1. Формулирование нулевой и альтернативной гипотез.

H_0: $\sigma_{e_1}^2 = \sigma_{e_2}^2 = \ldots = \sigma_{e_n}^2 = \sigma_e^2 = \text{const}$, т. е. дисперсия остатков однородна, остатки являются гомоскедастичными.

H_A: $\sigma_{e_i}^2 \neq \text{const}\left(i = \overline{1;n}\right)$, т. е. дисперсия остатков изменяется, остатки являются гетероскедастичными.

Шаг 2. Упорядочение наблюдений по росту элементов фактора.

Упорядочение данных по возрастанию фактора проводим в следующем порядке:

- выделим мышью весь массив исходных данных, т. е. **C3:D20** (см. рисунок 4.1);
- выберем в меню **Данные** команду **Сортировка**.

На экране появится окно диалога **Сортировка диапазона**. Зададим параметры окна диалога.

В поле **Сортировать по** выберем фактор *X*. Если исследуется многофакторная модель, выберите фактор, который, на ваш взгляд, обуславливает гетероскедастичность остатков.

Установим флажок **по возрастанию** и нажмем кнопку **OK**.

Примечание – В исследуемой задаче исходные данные упорядочены.

Шаг 3. Отбрасывание наблюдений с, содержащихся в середине вектора наблюдений.

Количество наблюдений, которые нужно отбросить, рассчитываем по формуле

$$c = \frac{4}{15} \cdot n = \frac{4}{15} \cdot 18 = 4{,}8$$

в ячейке **F21** (см. рисунок 4.1). Принимаем $c = 4$ и отбрасываем четыре наблюдения с середины векторов исходных данных, т. е. имеем массив **C10: D13**.

Шаг 4. Построение эконометрических моделей по методу 1МНК для двух полученных совокупностей наблюдений.

Построим две эконометрические модели по двум полученным совокупностями (от первого до седьмого наблюдения включительно и от двенадцатого до восемнадцатого наблюдения) методом 1МНК. Условия $n_1, n_2 > m+1$ выполняются, поскольку $n_1 = n_2 = 7$ и $7 > 1+1 = 2$. Для этого дважды используем в меню **Сервис** команду **Анализ данных**, инструмент анализа **Регрессия**:

- для построения первой модели вводим адрес диапазона зависимых переменных **C3:C9** и адрес диапазона независимых переменных **D3:D9**; а также имя нового листа «Модель 1»;

- для построения второй модели введем адрес диапазона зависимых переменных **C14:C20** и адрес диапазона независимых переменных **D14:D20**, а также имя нового листа «Модель 2».

Скопируем последовательно с листов «Модель 1» и «Модель 2» значения ячеек столбца *Коэффициенты* строк ***Y*-пересечение** и **Переменная *X*1** в ячейки **B29:C31** и **B33:C35** соответственно (см. рисунок 4.2). Таким образом, по двум совокупностям наблюдений получили две эконометрические модели:

$$\hat{y}_1 = 1{,}474681 + 0{,}045532x \text{ и } \hat{y}_2 = -0{,}25955 + 0{,}0403x.$$

Шаг 5. Вычисление статистики F.

Поскольку $n_1 = n_2$, то значения статистики F вычислим по формуле (4.6) в ячейке **C37** (рисунок 4.5). Для контроля вычислим эти значения и по формуле (4.7) в ячейке **E37**. Значения совпали и равны. $F \approx 11{,}05$.

	B	C	D	E	F
36					
37	$F=\frac{SSE_2}{SSE_1}$	11,0489169	$F=\frac{MSE_2}{MSE_1}$	11,0489169	
38	P=0,95	k1=n1-m-1	5		
39		k2=n2-m=1	5		
40	F(5;5)	5,05			
41	Вывод:	Присутствует гетероскедастичность остатков			
42					

а

	B	C	D	E
35				
36				
37	$F=\frac{SSE_2}{SSE_1}$	=Модель2!C13 / Модель1!C13	$F=\frac{MSE_2}{MSE_1}$	=Модель2!D13/Модель1!D13
38	P=0,95	k1=n1-m-1	=7-1-1	
39		k2=n2-m=1	=7-1-1	
40	F(5;5)	5,05		
41	Вывод:	=ЕСЛИ(C37<C40;"Гетероскедастичность остатков отсутствует"; "Присутствует гетероскедастичность остатков")		

б

а – значения; б – формулы

Рисунок 4.5 – Тест Гольдфельда – Квандта

Шаг 6. Определение по статистическим таблицам F-распределения Фишера критического значения F-критерия.

Критическое значение F-критерия находят по статистическим таблицам F-распределения Фишера по соответствующим значениям:

- доверительной вероятности $P = 0{,}95$;
- степеней свободы $k_1 = n_1 - m - 1 = 5$ и $k_2 = n_2 - m - 1 = 5$.

Полученное табличное значение F-критерия $F_{\text{кр}} = 5{,}05$ вводим в ячейку C40.

Критическое значение F-критерия можно вычислить, используя встроенную статистическую функцию **FРАСПРОБР** (0,05, 5, 5).

Шаг 7. Сравнение расчетного значения отношения F с критическим и интерпретация результатов.

Вывод о наличии гетероскедастичности в строке остатков делаем в ячейке **C41**, используя встроенную логическую функцию **ЕСЛИ** (см. рисунок 4.5). Поскольку $F > F_{кр}$, то отвергаем нулевую гипотезу, т. е. с надежностью $P = 0{,}95$ можно считать, что существует гетероскедастичность остатков.

V Оценка параметров модели ОМНК (методом Эйткена).

Поскольку имеет место гетероскедастичность остатков, то оценку параметров модели осуществим методом Эйткена (ОМНК). Оператор оценки имеет вид $\hat{A} = \left(X^T S^{-1} X\right)^{-1} X^T S^{-1} Y$.

1 Определение матриц, входящих в оператор Эйткена.

Сначала составим матрицу в диапазоне **J3:K20** (рисунок 4.6). Транспонированную матрицу X^T найдем в диапазоне **B43:S44** (рисунок 4.7), используя встроенную функцию **ТРАНСП**, т. е. = **ТРАНСП**(**J3:K20**).

Обратную матрицу

$$S^{-1} = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & \lambda_{18} \end{pmatrix}$$

определим в диапазоне **B46:S63** (рисунок 4.8), пользуясь гипотезой $\lambda_i = \dfrac{1}{x_i} \quad (i = \overline{1;18})$, в следующем порядке:

- для вычисления диагональных элементов матрицы в ячейку **B46** вводим формулу =1/B44, затем в ячейку **C47** вводим формулу =1/C44 и т. д.;

- в последние ячейки матрицы S^{-1} (вне главной диагонали) вводим нули.

	J	K
3	1	15
4	1	15
5	1	16
6	1	17
7	1	17
8	1	18
9	1	19
10	1	20
11	1	20
12	1	22
13	1	64
14	1	68
15	1	72
16	1	80
17	1	85
18	1	90
19	1	95
20	1	100

Рисунок 4.6 – Матрица X

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
43	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
44	15	15	16	17	17	18	19	20	20	22	64	68	72	80	85

Рисунок 4.7 – Транспонированная матрица X^T

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
46	0,07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	0	0,07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	0	0	0,06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0,06	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0,059	0	0	0	0	0	0	0	0	0	0	0	0	0
51	0	0	0	0	0	0,056	0	0	0	0	0	0	0	0	0	0	0	0
52	0	0	0	0	0	0	0,053	0	0	0	0	0	0	0	0	0	0	0
53	0	0	0	0	0	0	0	0,05	0	0	0	0	0	0	0	0	0	0
54	0	0	0	0	0	0	0	0	0,05	0	0	0	0	0	0	0	0	0
55	0	0	0	0	0	0	0	0	0	0,05	0	0	0	0	0	0	0	0
56	0	0	0	0	0	0	0	0	0	0	0,016	0	0	0	0	0	0	0
57	0	0	0	0	0	0	0	0	0	0	0	0,015	0	0	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0,014	0	0	0	0	0
59	0	0	0	0	0	0	0	0	0	0	0	0	0	0,013	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,012	0	0	0
61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,011	0	0
62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,011	0
63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,01

Рисунок 4.8 – Обратная матрица S^{-1}

2 Вычисление произведений матриц.

Произведение матриц $X^T S^{-1}$ вычислим, используя встроенную математическую функцию **=МУМНОЖ(B43:S44;B46:S63)**, в диапазоне **B65:S66** (рисунок 4.9). Произведение матриц $X^T S^{-1} X$ вычислим, используя встроенную математическую функцию **МУМНОЖ (B65:S66; J3:K20)**, в диапазоне **C68:D69** (рисунок 4.10).

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
65	0,0667	0,07	0,06	0,06	0,059	0,056	0,053	0,05	0,05	0,05	0,016	0,015	0,014	0,013	0,012	0,011	0,011	0,01
66	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Рисунок 4.9 – Произведение матриц $X^T S^{-1}$

3 Определение обратной матрицы $\left(X^T S^{-1} X\right)^{-1}$ и вектора $X^T S^{-1} Y$.

Обратную матрицу $\left(X^T S^{-1} X\right)^{-1}$ вычислим, используя встроенную математическую функцию **=МОБР(C68:D69)**, в диапазоне **F68:G69** (см. рисунок 4.10), а вектор $X^T S^{-1} Y$, используя встроенную математическую функцию **=МУМНОЖ(B65:S66;C3:C20)**, в диапазоне **I68:I69** (см. рисунок 4.10).

4 Вычисление вектора оценок параметров модели.

Вектор оценок параметров модели вычислим, используя встроенную математическую функцию =МУМНОЖ(F68:G69;I68:I69), в диапазоне K68:K69 (см. рисунок 4.10).

	B	C	D	E	F	G	H	I	J	K
68	$X^T S^{-1} X$	0,6672	18	$\left(X^T S^{-1} X\right)^{-1}$	3,5934	-0,0776	$X^T S^{-1} Y$	1,5981	Вектор	2,02230
69		18	833		-0,0776	0,0029		47,91	оцінок $\hat{A}$	0,01382

Рисунок 4.10 – Определение матриц $X^T S^{-1} X$, $\left(X^T S^{-1} X\right)^{-1}$ и векторов $X^T S^{-1} Y$ и $\hat{A}$

Таким образом, эконометрическая модель сбережений жителей города, которая оценена по ОМНК (методом Эйткена), имеет вид

$$\hat{y} = 2{,}02230 + 0{,}01382x.$$

VI Сравнительный анализ оценок параметров, полученных разными методами, выводы о смещенности оценок параметров моделей.

Оценим стандартные ошибки оценок параметров и сделаем выводы о смещенности этих оценок для каждой из двух полученных моделей.

Для модели, которая оценена по 1МНК

Скопируем значения стандартных ошибок из листа 1МНК (столбцы **Коэффициенты** и **Стандартная ошибка**) в ячейки **B87:D89** (рисунок 4.11). В ячейках **D88** и **D89** получили значение ошибок: $\hat{\sigma}_{a_0} = 0{,}13$, $\hat{\sigma}_{a_1} = 0{,}002$.

Сравним каждую стандартную ошибку с соответствующим значением оценки параметра с помощью формул

$$\frac{\hat{\sigma}_{\hat{a}_j}}{\left|\hat{a}_j\right|}, \quad j = \overline{0;1},$$

которые введем в ячейки **F88:F89**. Выводы о смещенности (несмещенности) оценок параметров модели сделаем, используя встроенную логическую функцию **ЕСЛИ** в ячейках **G88:G89**. Получили выводы: оценка параметра a_0 является несмещенной; оценка параметра a_1 является смещенной.

	B	C	D	E	F	G
86	1МНК	Дисперсионный анализ		Выводы о смещенности оценок параметров		
87		*Коэффициенты*	*Стандартная ошибка*			
88	Y-пересечение	2,000019	0,129962	$\frac{\hat{\sigma}_{\hat{a}_0}}{\lvert\hat{a}_0\rvert}$	6,4981	Оценка несмещена
89	Доход	0,014297	0,002291	$\frac{\hat{\sigma}_{\hat{a}_1}}{\lvert\hat{a}_1\rvert}$	16,026	Оценка смещена
90	ОМНК	Стандартная ошибка		Выводы о смещенности оценок параметров		
91	$\hat{\sigma}_{\hat{a}_0} = \sqrt{\hat{\sigma}^2_{\hat{a}_0}}$		0,077218	$\frac{\hat{\sigma}_{\hat{a}_0}}{\lvert\hat{a}_0\rvert}$	3,8183	Оценка несмещена
92	$\hat{\sigma}_{\hat{a}_1} = \sqrt{\hat{\sigma}^2_{\hat{a}_1}}$		0,002185	$\frac{\hat{\sigma}_{\hat{a}_1}}{\lvert\hat{a}_1\rvert}$	15,818	Оценка смещена

а

	B	C	D	E	F	G
86	1МНК	Дисперсионный анализ		Выводы о смещенности оценок параметров		
87		*='1МНК'!B16*	*='1МНК'!C16*			
88	='1МНК'!A17	='1МНК'!B17	='1МНК'!C17	$\frac{\hat{\sigma}_{\hat{a}_0}}{\lvert\hat{a}_0\rvert}$	=D88/C88*100	=ЕСЛИ(F88 >10;"Оценка смеще
89	='1МНК'!A18	='1МНК'!B18	='1МНК'!C18	$\frac{\hat{\sigma}_{\hat{a}_1}}{\lvert\hat{a}_1\rvert}$	=D89/C89*100	=ЕСЛИ(F89 >10;"Оценка смеще
90	ОМНК	Стандартная ошибка		Выводы о смещенности оценок параметров		
91		$\hat{\sigma}_{\hat{a}_0} = \sqrt{\hat{\sigma}^2_{\hat{a}_0}}$	=E83^0,5	$\frac{\hat{\sigma}_{\hat{a}_0}}{\lvert\hat{a}_0\rvert}$	=D91/K68*100	=ЕСЛИ(F91 >10;"Оценка смеще
92		$\hat{\sigma}_{\hat{a}_1} = \sqrt{\hat{\sigma}^2_{\hat{a}_1}}$	=F84^0,5	$\frac{\hat{\sigma}_{\hat{a}_1}}{\lvert\hat{a}_1\rvert}$	=D92/K69*100	=ЕСЛИ(F92 >10;"Оценка смеще

б

а – значения; б – формулы

Рисунок 4.11 – Выводы о смещенности оценок параметров

Для модели, которая оценена по ОМНК

Для вычисления стандартных ошибок прежде всего необходимо оценить дисперсию остатков и ковариационную матрицу оценок параметров модели по формулам, приведеным в таблице 4.1. Расчет проведем в следующем порядке:

- транспонированный вектор остатков e^T найдем в диапазоне **B73:S73** (рисунок 4.12), используя встроенную функцию **ТРАНСП**, т. е. **=ТРАНСП(Н3: Н20)**;

- произведение вектора на обратную матрицу $e^T \cdot S^{-1}$ найдем в диапазоне **B75:S75** (см. рисунок 4.12), используя встроенную функцию **МУМНІОЖ**, т. е. **= МУМНОЖ(B73:S73; B46:S63)**;

- произведение векторов $e^T \cdot S^1 \cdot e$ найдем в ячейке **F76** (см. рисунок 4.12), используя встроенную функцию = **МУМНОЖ(B75:S75;H3:H20)**;

- в ячейку **E81** (рисунок 4.13) введем формулу оценки дисперсий остатков $\hat{\sigma}_e^2 = \dfrac{e^T \cdot S^{-1} \cdot e}{n-m-1}$;

- для сравнения с аналогичной характеристикой модели, оцененной по 1МНК, в ячейку **C81** скопируем значения ячейки **Остаток** таблицы **Дисперсионный анализ**;

- в диапазоне **E83:F84** по формуле $\hat{\Sigma}_A = \hat{\sigma}_e^2 \cdot \left(X^T \cdot S^{-1} \cdot X\right)^{-1}$ определим элементы ковариационной матрицы;

- в ячейках **D91:D92** вычислим значения стандартных ошибок оценок параметров модели, как квадратные корни из диагональных элементов ковариационной матрицы (см. рисунок 4.11);

- в ячейках **G91:G92** сделаем выводы о смещенности оценок параметров модели таким же образом, как для предыдущей модели.

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
72	Транспонированный вектор остатков e^T																	
73	0,0705	-0,03	-0,163	-0,057	-0,1572	0,049	0,1652	0,201	-0,0986	0,1738	0,193	-0,562	-0,197	-0,08757	-0,5	0,644	-0,235	0,5861
74	Произведение $e^T \cdot S^{-1}$																	
75	0,0047	-0,002	-0,01	-0,003	-0,009	0,003	0,009	0,010	-0,005	0,008	0,003	-0,008	-0,003	-0,00109	-0,01	0,007	-0,002	0,006
76	Произвед $e^T \cdot S^{-1} \cdot e$				0,02655													

Рисунок 4.12 – Определение произведения $e^T \cdot S^{-1} \cdot e$

	B	C	D	E	F
78	Дисперсия остатков $\hat{\sigma}_e^2$				
79	1МНК	Дисперсионный анализ	ОМНК	По формуле	
80		MS			
81	Остаток	0,101631506		0,00166	
82	Ковариационная матрица			$\hat{\Sigma}_A = \hat{\sigma}_e^2 \cdot \left(X^T \cdot S^{-1} \cdot X\right)^{-1}$	
83		3,593394557	-0,07765	**0,005963**	-0,0001
84	0,001659	-0,077648382	0,00288	-0,00013	**5E-06**

а

	B	C	D	E	F
78	Дисперсия остатков $\hat{\sigma}_e^2$				
79	**1МНК**	**Дисперсионный анализ**	ОМНК:	По формуле	
80		MS			
81	Остаток	='1МНК'!D13		=F76/16	
82	Ковариационная матрица			$\hat{\Sigma}_A = \hat{\sigma}_e^2 \cdot \left(X^T \cdot S^{-1} \cdot X\right)^{-1}$	
83		=F68:G69	=F68:G69	**=$E$81*C83**	=$E$81*D83
84	=E81	=F68:G69	=F68:G69	=$E$81*C84	**=$E$81*D84**

б

а – значения; б – формулы

Рисунок 4.13 – Определение ковариационной матрицы

VII Проверка существенности связи, описываемая моделями, параметры которых оценены по методу Эйткена и методом 1МНК.

С целью проверки существенности связи, которая описывается моделями, вычислим коэффициенты множественной детерминации.

Для модели, которая оценена по ОМНК в ячейке **C97** (рисунок 4.14), вычислим коэффициент множественной детерминации по соответствующей формуле, приведенной в таблице 4.1.

	B	C	D	E
94	ОМНК		1МНК	
95	$Y^T \cdot Y$	133,1035	*Регрессионная статистика*	
96	$\bar{y}$	2,661666667	Множественный R	0,841869
97	R^2	0,995245	R-квадрат	0,708743
98	R	0,997620		

а

	B	C	D	E
94	ОМНК		1МНК	
95	$Y^T \cdot Y$	=СУММКВ(C3:C20)	='1МНК'!A3	
96	$\bar{y}$	=СРЗНАЧ(C3:C20)	='1МНК'!A4	='1МНК'!B4
97	R^2	=1-F76/(C95-18*C96^2)	='1МНК'!A5	='1МНК'!B5
98	R	=C97^0,5		

б

а – значения; б – формулы

Рисунок 4.14 – Оценка адекватности моделей

Для оценки адекватности модели, которая оценена по 1МНК, воспользуемся *регрессионной статисткой* итогового листа пакета **Анализ данных**.

Для модели, параметры которой оценены по ОМНК, $R^2 = 0{,}9952$ и $R = 0{,}9976$, а для модели, параметры которой оценены по 1МНК, $R^2 = 0{,}7087$ и $R = 0{,}8419$. Сравнительный анализ полученных значений и приводит к выводу, что ОМНК дает возможность получить адекватную эконометрическую модель. Для экономического анализа и прогнозирования сбережений жителей города следует использовать модель, которая оценена по ОМНК.

4.5 Варианты заданий

По статистическим данным, которые приведены в таблицах-вариантах 1–20 необходимо проверить гипотезу о наличии гетероскедастичности, построить эконометрические модели по методу Эйткена (ОМНК) и одношаговым методом наименьших квадратов (1МНК), сравнить модели и сделать выводы о точности оценок.

Вариант 1

Номер склада	Издержки на реализацию продукции, млн усл. ден. ед.	Объем товаро-оборота, млн т	Средний уровень товарных запасов, млн т
1	300	25	5
2	280	20	4
3	350	30	6
4	340	30	7
5	330	28	7
6	320	28	5
7	310	25	6
8	301	24	4
9	320	27	5
10	280	22	4
11	340	35	6
12	360	30	7
13	320	29	7
14	300	28	5
15	310	25	6

Вариант 2

Номер склада	Издержки на реализацию продукции, млн усл. ден. ед.	Объем товаро-оборота, млн т	Средний уро-вень товарных запасов, млн т
1	350	26	5
2	280	22	4
3	350	30	6
4	340	30	7
5	300	29	7
6	320	28	5
7	320	25	6
8	280	24	4
9	300	23	6
10	380	21	5
11	340	30	6
12	360	32	7
13	330	28	8
14	320	29	5
15	340	25	6

Вариант 3

Номер склада	Издержки на реализацию продукции, млн усл. ден. ед.	Объем товаро-оборота, млн т	Средний уровень товарных запасов, млн т
1	300	25	5
2	280	20	4
3	350	20	6
4	340	10	7
5	380	28	7
6	320	28	5
7	310	15	6
8	400	24	4
9	350	25	5
10	280	20	4
11	350	20	6
12	340	20	7
13	380	28	7
14	320	28	5
15	310	15	6
16	320	14	4

Вариант 4

Номер склада	Издержки на реализацию продукции, млн усл. ден. ед.	Объем товаро-оборота, млн т	Средний уро-вень товарных запасов, млн т
1	200	15	5
2	180	20	4
3	250	30	6
4	240	30	7
5	230	38	7
6	220	38	5
7	210	25	6
8	250	34	4
9	200	15	5
10	190	25	4
11	250	30	6
12	240	30	7
13	230	18	7
14	28	28	5
15	210	25	6
16	200	14	4

Вариант 5

Номер склада	Издержки на реализацию продукции, млн усл. ден. ед.	Объем товарооборота, млн т	Средний уровень товарных запасов, млн т
1	300	8	5
2	280	10	4
3	350	20	6
4	340	15	7
5	330	18	7
6	320	18	5
7	310	15	6
8	300	14	4
9	300	9	5
10	280	10	4
11	350	15	6
12	340	20	7
13	330	21	7
14	320	15	5
15	310	17	6
16	300	20	4

Вариант 6

Номер склада	Объем потребления на душу населения, т	Цена единицы продукции, усл. ден. ед.	Доход на душу населения, усл. ден. ед.
1	100	30	90
2	120	3 5	75
3	130	40	85
4	125	30	90
5	140	45	105
6	150	30	93
7	155	55	97
8	160	60	100
9	100	30	80
10	120	35	75
11	130	40	85
12	125	40	100
13	140	45	95
14	150	50	93
15	155	55	97
16	160	60	100

Вариант 7

Номер наблюдения	Объем потребления на душу населения, т	Цена единицы продукции, усл. ден. ед.	Доход на душу населения, усл. ден. ед.
1	100	30	8
2	120	35	17
3	130	40	18
4	125	30	19
5	140	45	19
6	150	50	19
7	155	55	19
8	160	60	20
9	100	30	9
10	120	35	7
11	130	40	28
12	125	30	29
13	140	45	9
14	150	50	9
15	155	55	19
16	160	60	30

Вариант 8

Номер наблюдения	Объем потребления на душу населения, т	Цена единицы продукции, усл. ден. ед.	Доход на душу населения, усл. ден. ед.
1	100	30	28
2	120	35	27
3	130	40	25
4	125	30	20
5	140	45	35
6	150	50	40
7	155	55	45
8	160	60	50
9	150	65	120
10	140	70	125
11	130	35	30
12	180	45	145
13	120	40	50
14	135	55	55
15	155	50	88
16	160	40	68

Вариант 9

Номер наблюдения	Потребление продукта на душу населения, т	Цена единицы продукции, усл. ден. ед.	Издержки на производство 1 т продукции, тыс. усл. ден. ед.
1	50	10	60
2	45	12	62
3	55	9	65
4	60	10	60
5	60	8	55
6	70	16	50
7	52	10	70
8	47	12	62
9	55	9	02
10	50	10	60
11	65	18	83
12	70	9	50

Вариант 10

Номер наблюдения	Потребление продукта на душу населения, т	Цена единицы продукции, усл. ден. ед.	Издержки на производство 1 т продукции, тыс. усл. ден. ед.
1	50	10	60
2	45	12	62
3	55	9	62
4	50	10	60
5	60	8	55
6	70	6	50
7	55	5	20
8	57	6	22
9	56	7	20
10	59	8	25
11	51	9	20
12	58	10	15

Вариант 11

Номер наблюдения	Потребление продукта на душу населения, т	Цена единицы продукции, усл. ден. ед.	Издержки на производство 1 т продукции, тыс. усл. ден. ед.
1	50	10	70
2	45	12	62
3	55	9	72
4	50	10	60
5	60	I8	55
6	70	6	60
7	50	15	60
8	45	12	82
9	55	19	62
10	50	10	60
11	60	9	85
12	70	14	60

Вариант 12

Номер наблюдения	Потребление продукта на душу населения, т	Цена единицы продукции, усл. ден. ед.	Издержки на производство 1 т продукции, тыс. усл. ден. ед.
1	50	10	90
2	45	12	62
3	55	19	82
4	50	10	60
5	60	8	85
6	70	16	50
7	50	10	70
8	45	12	62
9	55	9	62
10	50	10	80
11	60	12	65
12	70	14	60

Вариант 13

Номер наблюдения	Потребление продукта на душу населения, т	Цена единицы продукции, усл. ден. ед.	Издержки на производство 1 т продукции, тыс. усл. ден. ед.
1	50	10	70
2	45	12	62
3	55	9	62
4	50	10	60
5	60	8	55
6	70	6	50
7	80	8	82
8	85	10	65
9	75	12	74
10	90	7	87
11	45	15	62
12	50	14	60

Вариант 14

Номер наблюдения	Потребление продукта на душу населения, т	Цена единицы продукции, усл. ден. ед.	Издержки на производство 1 т продукции, тыс. усл. ден. ед.
1	55	5	100
2	57	6	110
3	56	7	112
4	59	8	115
5	51	9	120
6	58	10	135
7	60	8	140
8	50	9	120
9	45	15	110
10	65	8	130
11	70	11	150
12	52	12	170

Вариант 15

Номер наблюдения	Потребление продукта на душу населения, т	Цена единицы продукции, усл. ден. ед.	Издержки на производство 1 т продукции, тыс. усл. ден. ед.
1	75	5	100
2	67	6	110
3	56	7	112
4	59	8	115
5	51	9	120
6	58	10	135
7	60	7	120
8	65	8	130
9	45	9’	110
10	40	10	100
11	41	15	130
12	65	14	140
13	80	7	150
14	82	5	170
15	58	10	100

Вариант 16

Номер наблюдения	Потребление продукта на душу населения, т	Цена единицы продукции, усл. ден. ед.	Издержки на производство 1 т продукции, тыс. усл. ден. ед.
1	55	5	160
2	87	6	170
3	56	7	112
4	59	8	1 15
5	51	9	120
6	58	10	135
7	60	11	145
8	65	5	100
9	70	6	110
10	50	4	120
11	55	5	130
12	62	7	150
13	53	8	140
14	50	9	120
15	55	10	130

Вариант 17

Номер предпр.	Уровень рентабель-ности, %	Уровень издержек оборота, %	Произво-дитель-ность тру-да, усл. ден. ед.
1	7,89	8,9	17646
2	14,40	4,3	10177
3	10,11	4,9	14331
4	8,13	5,6	17010
5	8,85	6,2	15450
6	6,78	8,3	18172
7	8,91	7,8	17477
8	9,01	4,7	13137
9	6,13	10,8	19378
10	6,01	10,2	19343
11	6,17	13,1	22110
12	5,98	13,3	24111
13	6,10	10,7	19393
14	5,90	13,7	25445
15	6,00	11,1	21100

Вариант 18

Номер предпр.	Уровень рентабель-ности, %	Уровень издер-жек оборота, %	Производи-тельность труда, усл. ден. ед.
1	5,22	14,91	7134
2	4,41	15,05	5414
3	5,32	17,40	7633
4	6,72	11,55	8736
5	7,14	9,21	5590
6	4,40	14,20	7700
7	3,78	16,23	5923
8	6,83	11,97	6200
9	6,07	13,05	10212
10	6,10	13,45	10259
11	7,10	10,13	14620
12	6,21	12,33	12501
13	3,70	15,23	12255
14	5,55	13,95	11586
15	6,90	10,17	13450

Вариант 19

Номер предпр.	Уровень рентабель-ности, %	Уровень издержек оборота, %	Произво-дитель-ность тру-да, усл. ден. ед.
1	6,23	19,91	7234
2	5,41	20,05	5514
3	6,32	22,40	7733
4	7,72	16,55	8836
5	8,14	14,21	5690
6	5,40	19,20	7800
7	4,78	21,23	6023
8	7,83	16,97	6300
9	7,07	18,05	11212
10	7,10	18,45	11259
11	8,10	15,13	15620
12	7,21	7,33	13501
13	4,70	20,23	13255
14	6,55	18,95	12586
15	7,90	15,17	14450

Вариант 20

Номер предпр.	Уровень рентабель-ности, %	Уровень издержек оборота, %	Производи-тельность труда, усл. ден. ед.
1	12,89	13,9	27649
2	19,40	9,3	20170
3	15,11	10,6	20170
4	13,13	10,6	27013
5	13,85	11,2	25453
6	11,78	13,3	28175
7	13,91	12,8	27470
8	14,01	9,7	23130
9	11,13	15,8	29371
10	11,01	15,2	29346
11	11,17	18,1	32113
12	10,98	18,3	34114
13	12,10	15,7	29396
14	10,90	18,7	35448
15	12,00	16,1	31103

Практическая работа №5
Мультиколлинеарность

5.1 Понятие мультиколлинеарности. Алгоритм Феррара – Глобера

Одним из предположений применения одношагового метода наименьших квадратов (1МНК) является предположение, которое касается матрицы исходных данных *X*: факторы модели должны быть независимыми между собой, т. е. отсутствие мультиколлинеарности [1].

Мультиколлинеарность означает существование тесной линейной зависимости или сильной корреляции между двумя или более факторами.

Мультиколлинеарность негативно влияет на количественные характеристики эконометрической модели.

Основные следствия мультиколлинеарности:

- большие дисперсии и ковариации оценщиков параметров модели и, как следствие, проблематичное тестирование и возвеличивание интервалов доверия для параметров и зависимой переменной;

- невозможность получить оценки параметров модели 1МНК при экстремальной мультиколлинеарности, потому что в этом случае определитель $\left|X^T \cdot X\right| = 0$ и обратная матрица $(X^T \cdot X)^{-1}$ не существуют;

- нестабильность оценок параметров модели. Оценки параметров становятся чувствительными к изменениям объемов совокупности наблюдений: небольшие изменения в массиве данных являются причиной больших изменений в оценках параметров.

Поэтому в эконометрических исследованиях важно выявить: существуют ли взаимосвязи между факторами, которые являются мультиколлинеарностью. Существует много методов исследования мультиколлинеарности, но единой меры пока не предложено. Большинство этих методов заключается в исследовании корреляционной матрицы.

В эконометрических задачах для исследования наличия мультиколлинеарности широко применяется метод Феррара – Глобера [7]. Этот метод включает три статистических критерия, на основании которых проверяется гипотеза об отсутствии мультиколлинеарности:

- в массиве факторов (критерий χ^2);

- каждого фактора со всеми другими факторами массива (*F*-критерий Фишера);

- между двумя факторами (*t*-критерий Стьюдента).

Все эти критерии при сравнении с их критическим значением дают возможность сделать конкретные выводы о наличии или отсутствии мультиколлинеарности факторов с определенным уровнем значимости.

Рассмотрим алгоритм Феррара – Глобера.

Шаг 1. Стандартизация переменных.

Шаг 2. Нахождение корреляционной матрицы *R*.

Шаг 3. Критерий χ^2.

Шаг 4. Определение матрицы, обратной матрице *R*.

Шаг 5. *F*-критерий Фишера.

Шаг 6. Расчет коэффициентов частной корреляции.

Шаг 7. *T*-критерий Стьюдента.

Общий недостаток всех мер и методов заключается в том, что ни один из них не проводит четкой грани между тем, когда мультиколлинеарность можно считать существенной, и тем, когда мультиколлинеарность можно проигнорировать. Вторая проблема состоит в том, как избавиться от опасной для моделирования мультиколлинеарности. Самый простой способ устранения мультиколлинеарности в массиве факторов – это исключение из рассмотрения тех факторов, которые являются причиной мультиколлинеарности. Но при этом могут возникнуть новые трудности: во-первых, устранение некоторых факторов может отразиться на теоретико-экономическом содержании модели; во-вторых, устранение существенных факторов приводит к смешанности МНК-оценок. Второй способ заключается в замене двух мультиколлинеарных факторов одним, равным их отношению, т. е. $\frac{x_i}{x_j}$. Этот способ также ухудшает интерпретацию модели.

5.2 Проверка эконометрической модели на мультиколлинеарность в среде табличного процессора MS Excel

К числу основных показателей технической оснащенности строительного производства относятся: фондовооруженность, механовооруженность и энерговооруженность труда рабочих. Чтобы построить эконометрическую модель зависимости среднегодового объема строительно-монтажных работ от этих показателей на основе метода 1МНК, необходимо быть уверенным, что между факторами не существует мультиколлинеарности. Исследовать наличие мультиколлинеарности между этими факторами на основе данных по десяти строительно-монтажным управлениям объединения, которые приведены на рисунке 5.1.

	B	C	D	E	F	G	H	I	J	K
1										
2	№	X1	X2	X3	$x_{i1}-\bar{x}_1$	$x_{i2}-\bar{x}_2$	$x_{i3}-\bar{x}_3$	x^*_{i1}	x^*_{i2}	x^*_{i3}
3	1	32	0,59	15,5	3,1	0,004	0,4	0,2209	0,0091	0,0491
4	2	29	0,43	10,5	0,1	-0,156	-4,6	0,0071	-0,3531	-0,5645
5	3	30	0,7	16,5	1,1	0,114	1,4	0,0784	0,2580	0,1718
6	4	31	0,61	17,5	2,1	0,024	2,4	0,1497	0,0543	0,2945
7	5	25	0,51	12,5	-3,9	-0,076	-2,6	-0,2779	-0,1720	-0,3191
8	6	34	0,51	13,5	5,1	-0,076	-1,6	0,3635	-0,1720	-0,1964
9	7	29	0,65	17,5	0,1	0,064	2,4	0,0071	0,1448	0,2945
10	8	24	0,43	13,5	-4,9	-0,156	-1,6	-0,3492	-0,3531	-0,1964
11	9	20	0,51	14,5	-8,9	-0,076	-0,6	-0,6343	-0,1720	-0,0736
12	10	35	0,92	19,5	6,1	0,334	4,4	0,4347	0,7559	0,5400
13	СУММ	289	5,86	151				*Матрица стандартизи-рованных переменных, X**		
14	СРЗНАЧ	28,9	0,586	15,1						
15	СУММКВ				196,9	0,19524	66,4			

а

	B	C	D	E	F	G	H	I	J	K
1										
2	№	X1	X2	X3	$x_{i1}-\bar{x}_1$	$x_{i2}-\bar{x}_2$	$x_{i3}-\bar{x}_3$	x^*_{i1}	x^*_{i2}	x^*_{i3}
3	1	32	0,59	15,5	=C3-$C$14	=D3-$D$14	=E3-$E$14	=F3/$F$15^0,5	=G3/$G$15^0,5	=H3/$H$15^0,5
4	2	29	0,43	10,5	=C4-$C$14	=D4-$D$14	=E4-$E$14	=F4/$F$15^0,5	=G4/$G$15^0,5	=H4/$H$15^0,5
5	3	30	0,7	16,5	=C5-$C$14	=D5-$D$14	=E5-$E$14	=F5/$F$15^0,5	=G5/$G$15^0,5	=H5/$H$15^0,5
6	4	31	0,61	17,5	=C6-$C$14	=D6-$D$14	=E6-$E$14	=F6/$F$15^0,5	=G6/$G$15^0,5	=H6/$H$15^0,5
7	5	25	0,51	12,5	=C7-$C$14	=D7-$D$14	=E7-$E$14	=F7/$F$15^0,5	=G7/$G$15^0,5	=H7/$H$15^0,5
8	6	34	0,51	13,5	=C8-$C$14	=D8-$D$14	=E8-$E$14	=F8/$F$15^0,5	=G8/$G$15^0,5	=H8/$H$15^0,5
9	7	29	0,65	17,5	=C9-$C$14	=D9-$D$14	=E9-$E$14	=F9/$F$15^0,5	=G9/$G$15^0,5	=H9/$H$15^0,5
10	8	24	0,43	13,5	=C10-$C$14	=D10-$D$14	=E10-$E$14	=F10/$F$15^0,5	=G10/$G$15^0,5	=H10/$H$15^0,5
11	9	20	0,51	14,5	=C11-$C$14	=D11-$D$14	=E11-$E$14	=F11/$F$15^0,5	=G11/$G$15^0,5	=H11/$H$15^0,5
12	10	35	0,92	19,5	=C12-$C$14	=D12-$D$14	=E12-$E$14	=F12/$F$15^0,5	=G12/$G$15^0,5	=H12/$H$15^0,5
13	СУММ	=СУММ(C3:C12)	=СУММ(D3:D12)	=СУММ(E3:E12)				*Матрица стандартизированных переменных, X**		
14	СРЗНАЧ	=СРЗНАЧ(C3:C12)	=СРЗНАЧ(D3:D12)	=СРЗНАЧ(E3:E12)						
15	СУММКВ				=СУММКВ(F3:F12)	=СУММКВ(G3:G12)	=СУММКВ(H3:H12)			

б

а – значение; б – формулы

Рисунок 5.1 – Исходные данные, построение матрицы стандартизированных переменных

Поставленную задачу решим в среде Excel по следующему алгоритму.

1 Идентификация переменных:

Y – среднегодовой объем строительно-монтажных работ – результативный показатель;

X_1 – фондовооруженность – показатель-фактор;

X_2 – механовооруженности – показатель-фактор;

X_3 – энерговооруженность – показатель-фактор.

Блок данных формируется в массиве **C3:E12**, в ячейки **B3:B12** вводятся порядковые номера строительно-монтажных управлений (рисунок 5.1, а).

2 Исследование наличия мультиколлинеарности по алгоритму Феррара – Глобера.

Шаг 1. Стандартизация переменных.

Элементы стандартизированных векторов рассчитываются по формуле

$$x_{ij}^{*} = \frac{x_{ij} - \overline{x_j}}{\sqrt{n\sigma_{x_j}^{2}}},\ i = \overline{1;n}\ \ j = \overline{1;m}, \tag{5.1}$$

где n – количество наблюдений;

m – количество факторов;

$\sigma_{x_j}^{2}$ – дисперсия j-го фактора.

Поскольку дисперсия рассчитывается по формуле

$$\sigma_{x_j}^{2} = \frac{\sum\limits_{i-1}^{n}\left(x_{ij} - \overline{x_j}\right)^{2}}{n}, \tag{5.2}$$

формулы для стандартизированных переменных примут вид

$$x_{ij}^{*} = \frac{x_{ij} - \overline{x_j}}{\sqrt{\sum\limits_{i-1}^{n}\left(x_{ij} - \overline{x_j}\right)^{2}}},\ i = \overline{1;n},\ j = \overline{1;m}. \tag{5.3}$$

Исходные данные факторов, размещенные в блоке **C3:E12** (рисунок 5.1), стандартизируем по формулам (5.3), используя встроенные статистические функции **СУММ, СУММКВ, СРЕДН** или по соответствующим формулам. В результате вычислений в массиве **I3:K12** получим матрицу стандартизированных переменных X^*.

Шаг 2. Нахождение корреляционной матрицы R (матрицы моментов стандартизированной системы нормальных уравнений).

Корреляционная матрица R определяется по формуле

$$R = X^{*^T} \cdot X^*, \qquad (5.4)$$

где X^* – матрица стандартизированных переменных.

Для нахождения элементов корреляционной матрицы R последовательно используем встроенные функции **Транспонирование матриц** – **ТРАНСП** и **Произведение матриц** – **МУМНОЖ** (рисунок 5.2).

	B	C	D	E	F	G	H	I	J	K
17	*Транспонированная матрица стандартизированных переменных*									
18	0,2209	0,0071	0,0784	0,1497	-0,2779	0,3635	0,0071	-0,3492	-0,6343	0,4347
19	0,0091	-0,3531	0,2580	0,0543	-0,1720	-0,1720	0,1448	-0,3531	-0,1720	0,7559
20	0,0491	-0,5645	0,1718	0,2945	-0,3191	-0,1964	0,2945	-0,1964	-0,0736	0,5400
21	*Корреляционная матрица*									
22		1	0,5751	0,4338		Проверка		1	0,5751	0,4338
23	*R*=	0,5751	1	0,8815			*R*=	0,5751	1	0,8815
24		0,4338	0,8815	1				0,4338	0,8815	1

а

	B	C	D	E	F
17					*Транспонированная матрица*
18	=ТРАНСП(I3:K1	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K12)
19	=ТРАНСП(I3:K1	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K12)
20	=ТРАНСП(I3:K1	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K12)
21					*Корреляционная м*
22		=МУМНОЖ(B18:K2	=МУМНОЖ(B18:K2	=МУМНОЖ(B18:K2	
23	*R*=	=МУМНОЖ(B18:K2	=МУМНОЖ(B18:K2	=МУМНОЖ(B18:K2	
24		=МУМНОЖ(B18:K2	=МУМНОЖ(B18:K2	=МУМНОЖ(B18:K2	

	G	H	I	J	K
17	*стандартизированных переменных*				
18	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K1	=ТРАНСП(I3:K12	=ТРАНСП(I3:K12
19	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K1	=ТРАНСП(I3:K12	=ТРАНСП(I3:K12
20	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K12)	=ТРАНСП(I3:K1	=ТРАНСП(I3:K12	=ТРАНСП(I3:K12
21	*иатрица*				
22	Проверка		1	=КОРРЕЛ(C3:C1	=КОРРЕЛ(C3:C
23		*R*=	=J22	1	=КОРРЕЛ(D3:D
24			=K22	=K23	1

б

а – значение; б – формулы

Рисунок 5.2 – Нахождение корреляционной матрицы

Проверку вычислений (блок **И22:К24**) целесообразно выполнить, используя последовательно встроенную статистическую функцию **КОРРЕЛ**, учитывая при этом свойства корреляционной матрицы: корреляционная матрица симметрична; на главной диагонали расположены единицы.

В ячейке **J22** находим коэффициент корреляции между факторами X_1 и X_2, в ячейке **K22** – коэффициент корреляции между факторами X_1 и X_3, в ячейке **K23** – коэффициент корреляции между факторами X_2 и X_3. Последние элементы – по свойствам.

Вывод: на основе значения коэффициента корреляции $r_{x_2x_3} = 0{,}8815$ можно сделать предварительный вывод о наличии возможной мультиколлинеарности между факторами X_2 и X_3 (рисунок 5.2, а).

Шаг 3. Критерий – χ^2.

Расчетное значение критерия χ^2 определяется по формуле

$$\chi^2_{\text{расч}} = -\left(n - 1 - \frac{1}{6}(2m+5)\right)\ln|R|, \tag{5.5}$$

где $|R|$ – определитель корреляционной матрицы;

R – *детерминант корреляции.*

По принятой доверительной вероятности P и числу степеней свободы $k = \frac{1}{2} \cdot m(m-1)$ определяется критическое значение критерия $\chi^2_{\text{кр}}$, которое сравнивается с расчетным:

- если $\chi^2_{\text{расч}} < \chi^2_{\text{кр}}$, то нет оснований отклонить гипотезу об отсутствии мультиколлинеарности в массиве факторов, т. е. с принятой надежностью можно утверждать, что в массиве факторов мультиколлинеарность отсутствует;

- если $\chi^2_{\text{расч}} > \chi^2_{\text{кр}}$, то гипотеза об отсутствии мультиколлинеарности в массиве факторов отклоняется, т. е. с принятой надежностью можно утверждать, что в массиве факторов мультиколлинеарность существует.

Замечание – Если гипотеза об отсутствии мультиколлинеарности в массиве факторов принимается, то исследование мультиколлинеарности прекращается.

Исследование наличия мультиколлинеарности в массиве факторов χ^2 по критерию в оболочке электронных таблиц Excel:

1 Находим определитель матрицы в ячейке **C26**, используя встроенную математическую функцию **МОПРЕД** (рисунок 5.3).

2 Находим натуральный логарифм определителя в ячейке **C27**, используя встроенную математическую функцию **LN**.

3 Находим расчетное значение критерия в ячейке **C28** по формуле (5.5).

4 В ячейку **C29** вводим значение уровня значимости $\alpha = 0{,}05$.

5 В ячейку **C30** – число степеней свободы $k = m = 3$.

6 В ячейке **C31** вычисляем критическое значение критерия χ^2 с помощью встроенной статистической функции **ХИ2ОБР**.

7 Делаем вывод о наличии мультиколлинеарности в массиве факторов в ячейке **С32**, используя встроенную логическую функцию **ЕСЛИ**.

	B	C
26	Определитель коррел.матрицы	0,1438
27	Натуральный логарифм опред-ля	-1,9393
28	Расчетное значение критерия	13,8982
29	Уровень значимости	0,05
30	Число степеней свободы	3
31	Критич.значение критерия Хи-кв	7,8147
32	Выводы о наличии мультиколлинеар в массиве факторов	Гипотеза об отсутствии мультиколлинеар-ти в массиве факторов отклоняется, т.е. с принятой надежн. можно утверждать, что в массиве фак-в мульт.существует

а

	B	C
26	Определитель коррел.матрицы	=МОПРЕД(C22:E24)
27	Натуральный логарифм опред-ля	=LN(C26)
28	Расчетное значение критерия	=-(10-1-1/6*(2*3+5))*C27
29	Уровень значимости	0,05
30	Число степеней свободы	3
31	Критич.значение критерия Хи-кв	=ХИ2ОБР(C29;C30)
32	Выводы о наличии мульт-ти в массиве факторов	=ЕСЛИ(C28<C31;"Нет оснований отклонят

б

а – значение; б – формулы

Рисунок 5.3 – Критерий χ^2

Выводы:

- на основе значения детерминанта корреляции $|R| = 0{,}1481(\to 0)$ можно сделать предварительный вывод о наличии возможной мультиколлинеарности в массиве факторов;

- на основе критерия χ^2 с надежностью P = 0,95 можно утверждать, что в массиве факторов мультиколлинеарность существует.

Шаг 4. Определение матрицы, обратной матрице R:

$$C = R^{-1} = \left(X^{*T} \cdot X^{*} \right)^{-1}. \tag{5.6}$$

Находим обратную матрицу C в массиве **С34:Е36**, используя встроенную математическую функцию **МОБР** (рисунок 5.4).

Шаг 5. F-критерий Фишера.

Расчетные значения критерия для каждого фактора отличаются по формуле

$$F_j = (c_{jj} - 1)\frac{n-m}{m-1}, \quad j = 1,\ 2, \ldots, m, \tag{5.7}$$

где c_{jj} – диагональные элементы матрицы $C = R^{-1}$.

По принятой доверительной вероятности и числу степеней свободы: $k_1 = m - 1$ – степень свободы знаменателя; $k_2 = n - m$ – степень свободы числителя $(k_1 < k_2)$ – определяется критическое значение *F*-критерия, которое сравнивается с расчетным:

- если $F_{j\,\text{расч}} < F_{j\,\text{кр}}$, то нет оснований отклонить гипотезу об отсутствии мультиколлинеарности между j-м фактором и остальными факторами массива, т. е. с принятой надежностью можно утверждать, что между j-м фактором и другими факторами мультиколлинеарность отсутствует;

- если $F_{j\,\text{расч}} > F_{J\,\text{кр}}$, то гипотеза об отсутствии мультиколлинеарности между j-м фактором и остальными факторами массива отклоняется, т. е. с принятой надежностью можно утверждать, что между j-м фактором и другими факторами мультиколлинеарность существует (см. рисунок 5.4).

Исследование наличия мультиколлинеарности каждого фактора со всеми другими факторами массива по *F*-критерию Фишера в оболочке электронных таблиц Excel:

1 Находим в ячейке **C37, C38, C39** по формуле (5.7) расчетные значения критерия F_1, F_2, F_3 соответственно (рисунок 5.4).

2 В ячейку **C40** вводим значение уровня значимости $\alpha = 0{,}05$.

3 В ячейках **C41** и **D41** вычисляем число степеней свободы $k_1 = m - 1 = 3 - 1 = 2$; $k_2 = m - n = 10 - 3 = 7$.

4 В ячейке **C42** вычисляем критическое значение F-критерия с помощью встроенной статистической функции **FРАСПОБР**.

5 Делаем вывод о наличии мультиколлинеарности фактора X_1 с факторами X_2 и X_3 в ячейке **D37**, используя встроенную логическую функцию **ЕС-ЛИ**. Поскольку функция будет копироваться в остальные ячейки столбца, то при вводе адреса ячеек, которые сравниваются, надо использовать абсолютную и относительную ссылки.

6 Копируем полученную формулу в ячейки **D38** и **D39** и делаем выводы о наличии мультиколлинеарности фактора X_2 с факторами X_1 и X_3, фактора X_3 с факторами X_1 и X_2.

	B	C	D	E	F	G	H	I	J
34	*Обратная корреляционная матрица*	1,5500	-1,3403	0,5092					
35		-1,3403	5,6453	-4,3951					
36		0,5092	-4,3951	4,6536					
37	*Значения F1 и вывод*	1,9250	Нет оснований отклонять гипотезу об отсутствии мультиколлинеарности между фактором и остальными факторами массива						
38	*Значения F2 и вывод*	16,2586	Гипотеза об отсутствии мультиколлинеарности между данным фактором и остальными факторами массива отклоняется						
39	*Значения F3 и вывод*	12,7875	Гипотеза об отсутствии мультиколлинеарности между данным фактором и остальными факторами массива отклоняется						
40	*Уровень значимости*	0,05							
41	*Число степеней свободы*	2	7						
42	*Критич. значение F-критерия*	4,7374							

а

	B	C	D	E
34	*Обратная корреляцион матрица*	=МОБР(C22:E24)	=МОБР(C22:E24)	=МОБР(C22:E24)
35		=МОБР(C22:E24)	=МОБР(C22:E24)	=МОБР(C22:E24)
36		=МОБР(C22:E24)	=МОБР(C22:E24)	=МОБР(C22:E24)
37	*Значения F1 и вывод*	=(C34-1)*(10-3)/(3-1)	=ЕСЛИ(C37<$C$42;"Нет оснований отк	
38	*Значения F2 и вывод*	=(D35-1)*(10-3)/(3-1)	=ЕСЛИ(C38<$C$42;"Нет оснований отк	
39	*Значения F3 и вывод*	=(E36-1)*(10-3)/(3-1)	=ЕСЛИ(C39<$C$42;"Нет оснований отк	
40	*Уровень значимости*	0,05		
41	*Число степеней свободы*	=3-1	=10-3	
42	*Критич. значение F-критер*	=FРАСПОБР(C40;C41;D41)		

б

а – значение; б – формулы

Рисунок 5.4 – F-критерий Фишера

Выводы: с надежностью P = 0,95 можно считать, что:

- между фактором X_1 и факторами X_2 и X_3 не существует мультиколлинеарности;
- между фактором X_2 и факторами X_1 и X_3 существует мультиколлинеарность;

- между фактором X_3 и факторами X_2 и X_1 существует мультиколлинеарность.

Шаг 6. Расчет коэффициентов частной корреляции.

Коэффициенты частной корреляции рассчитываются по формуле

$$r_{kj} = \frac{-c_{kj}}{\sqrt{c_{kk} \cdot c_{jj}}}, \ k = \overline{1;m}, \ j = \overline{1;m}, \qquad (5.8)$$

где c_{jj}, c_{kk} – диагональные элементы матрицы $C = (R)^{-1}$;

c_{kj} – элемент матрицы $C = (R)^{-1}$, который находится в k-й строке j-го столбца.

Поскольку для массива факторов, который исследуется, $m = 3$, то необходимо рассчитать три коэффициента частной корреляции: $r_{12(3)}$, $r_{13(2)}$, $r_{23(1)}$.

Находим $r_{12(3)}$, $r_{13(2)}$, $r_{23(1)}$ в ячейках **С45, С46, С47** (рисунок 5.5) соответственно, используя формулу (5.8).

Шаг 7. Т-критерий Стьюдента.

Расчетные значения t-критерия для каждой пары факторов определяются по формуле

$$t_{kj} = \frac{r_{kj}\sqrt{n-m}}{\sqrt{1-r_{kj}^2}}, \ k = \overline{1;m}, \ j = \overline{1;m}, \qquad (5.9)$$

где r_{kj} – соответствующие коэффициенты частной корреляции.

По принятой доверительной вероятности P и числу степеней свободы $k = n - m$ определяется критическое значение t-критерия, которое сравнивается с расчетным:

- если $\left|t_{kj\ \text{расч}}\right| < t_{kj\ \text{кр}}$, то нет оснований отклонить гипотезу об отсутствии мультиколлинеарности между k-м и j-м факторами, т. е. с принятой надежностью можно утверждать, что между k-м и j-м факторами мультиколлинеарность отсутствует;

- если $\left|t_{kj\ \text{расч}}\right| > t_{kj\ \text{кр}}$, то гипотеза об отсутствии мультиколлинеарности между k-м и j-м факторами отклоняется, т. е. с принятой надежностью можно утверждать, что между k-м и j-м факторами мультиколлинеарность существует.

Исследование наличия мультиколлинеарности каждой пары факторов по t-критерию Стьюдента в оболочке электронных таблиц Excel:

1 Расчетное значение критерия находим по формуле (5.9) в ячейках **C49, C50, C51** (рисунок 5.5).

2 В ячейку **C52** вводим значение уровня значимости.

3 В ячейке **C53** вычисляем число степеней свободы $k = n - m = 10 - 3 = 7$.

4 В ячейке **C54** вычисляем критическое значение t-критерия с помощью встроенной статистической функции **СТЬЮДРАСПОБР**.

5 Модуль расчетного значения критерия находим в ячейке **D49**, используя встроенную математическую функцию **ABS**, при этом делаем относительную ссылку на столбец. Полученную формулу копируем в ячейки **D50, D51**.

6 Делаем вывод о наличии мультиколлинеарности между факторами и в ячейке **E49**, используя встроенную логическую функцию **ЕСЛИ**. При этом делаем абсолютную и относительную ссылку.

7 Полученную формулу копируем в ячейки **E50** и **E51** и делаем выводы о наличии мультиколлинеарности между факторами X_1 и X_3, X_2 и X_3.

Выводы: с надежностью P = 0,95 можно утверждать, что:

- между факторами X_1 и X_2 мультиколлинеарность отсутствует;
- между факторами X_1 и X_3 мультиколлинеарность отсутствует;
- между факторами X_2 и X_3 мультиколлинеарность существует.

Общий вывод. Таким образом, с целью устранения мультиколлинеарности необходимо один из факторов механовооруженности (X_2) или энерговооруженности (X_3) труда рабочих исключить из рассмотрения, если это целесообразно с экономической точки зрения и не влечет за собой устранение из модели существенного фактора.

	B	C	D	E	F	G	H	I	J	K
44	*Коэффициенты частичной*									
45	$r_{12(3)}$	0,4531								
46	$r_{13(2)}$	-0,1896								
47	$r_{23(1)}$	0,8575								
48	*Значение t-критерия*		*Модуль*	*Выводы о наличии мультиколлинеарности*						
49	$t_{12(3)}$	1,3448	1,3448	Нет оснований отклонять гипотезу об отсутствии мультиколлин-ти между факторами, т.е. с принятой надежностью можно утверждать, что между X1 и X2 мультиколлинеарность отсутствует						
50	$t_{13(2)}$	-0,5108	0,5108	Нет оснований отклонять гипотезу об отсутствии мультиколлин-ти между факторами, т.е. с принятой надежностью можно утверждать, что междуX3 и X1 мультиколлинеарность отсутствует						
51	$t_{23(1)}$	4,4097	4,4097	Гипотеза об отсутствии мультиколлинеарности между данными факторами отклоняется, т.е. с принятой надежностью можно утверждать, что между X2 и X3 мультик-ть существует						
52	*Уровень значимости*	0,05								
53	*Число степеней свободы*	7								
54	*Критич. значение t-критерия*	2,3646								

а

	B	C	D	E	F
44	*Коэффициенты частичной корреляции*				
45	$r_{12(3)}$	=-D34/(C34*D35)^0,5			
46	$r_{13(2)}$	=-E34/(C34*E36)^0,5			
47	$r_{23(1)}$	=-E35/(D35*E36)^0,5			
48	*Значение t-критерия*		*Модуль*		
49	$t_{12(3)}$	=C45*(10-3)^0,5/(1-C45^2)^0,5	=ABS(C49)	=ЕСЛИ(D49<$C$54;"Нет оснований от	
50	$t_{13(2)}$	=C46*(10-3)^0,5/(1-C46^2)^0,5	=ABS(C50)	=ЕСЛИ(D50<$C$54;"Нет оснований от	
51	$t_{23(1)}$	=C47*(10-3)^0,5/(1-C47^2)^0,5	=ABS(C51)	=ЕСЛИ(D51<$C$54;"Нет оснований от	
52	*Уровень значимости*	0,05			
53	*Число степеней свободы*	=10-3			
54	*Критич. значение F-критер*	=СТЬЮДРАСПОБР(C52;C53)			

б

а – значение; б – формулы

Рисунок 5.5 – *Т*-критерий Стьюдента

5.3 Варианты заданий

Предположим, что на уровень рентабельности предприятий общественного питания существенно влияют следующие показатели хозяйственной деятельности: относительный уровень издержек обращения (%), доля продукции собственного производства (%), численность работников в расчете на 1 тыс. усл. ден. ед. товарооборота (чел.).

Чтобы построить эконометрическую модель этой зависимости по методу 1МНК, необходимо быть уверенным, что между факторами относительного уровня издержек обращения, долей собственной продукции и трудоемкостью не существует мультиколлинеарности. Исследовать наличие мультиколлинеарности между этими факторами по данным десяти предприятий общественного питания города, которые приведены в таблицах-вариантах 1–24 (данные условные).

Вариант 1

№	Уровень расходов	Собственная продукция	Трудоемкость
1	15,6	40,4	21,1
2	13,5	38,9	27,8
3	15,3	36,6	21,7
4	14,9	41,4	21,5
5	15,1	32,2	21,1
6	16,1	31,4	19,7
7	16,7	32,6	19,6
8	15,4	38,7	21,2
9	17,1	44,3	20,2
10	16,8	39,3	21,3

Вариант 2

№	Уровень расходов	Собственная продукция	Трудоемкость
1	15,1	32,0	21,2
2	16,1	32,4	19,8
3	16,7	32,6	19,6
4	15,5	38,7	21,5
5	17,2	44,3	20,2
6	16,9	39,3	20,5
7	17,0	40,4	20,2
8	16,2	41,5	21,3
9	15,0	45,2	21,4
10	18,0	50,2	19,0

Вариант 3

№	Уровень расходов	Собственная продукция	Трудоемкость
1	16,9	39,2	20,2
2	16,2	41,0	21,3
3	15,5	41,3	31,4
4	18,2	45,2	18,9
5	17,3	50,2	24,8
6	17,1	51,6	19,4
7	16,4	48,0	19,3
8	16,7	48,6	19,6
9	14,2	49,8	25,7
10	17,2	45,0	22,1

Вариант 4

№	Уровень расходов	Собственная продукция	Трудоемкость
1	17,2	51,8	24,8
2	17,1	50,4	19,5
3	16,5	48,0	19,2
4	16,8	48,6	19,6
5	14,5	49,8	26,8
6	17,2	45,0	22,2
7	17,1	40,4	22,3
8	17,9	41,7	21,5
9	16,2	38,8	24,1
10	17,3	40,6	22,5

Вариант 5

№	Уровень расходов	Собственная продукция	Трудоемкость
1	14,3	49,8	26,8
2	17,2	45,2	22,1
3	17,0	41,4	22,4
4	17,8	41,7	21,5
5	16,3	38,7	24,1
6	17,3	40,6	22,5
7	16,9	33,6	21,3
8	14,8	13,9	25,6
9	19,6	32,5	19,0
10	11,4	52,5	30,0

Вариант 6

№	Уровень расходов	Собственная продукция	Трудоемкость
1	17,8	41,8	21,7
2	16,3	38,7	24,0
3	17,2	40,5	22,4
4	16,8	34,6	21,3
5	14,9	13,9	25,5
6	19,6	31,5	19,0
7	11,5	52,5	30,0
8	17,2	37,0	22,2
9	19,5	43,6	19,6
10	12,5	48,3	28,2

Вариант 7

№	Уровень расходов	Собственная продукция	Трудоемкость
1	16,9	33,7	21,3
2	14,5	13,8	25,6
3	19,5	31,5	19,1
4	11,5	52,5	30,0
5	17,1	37,0	22,3
6	19,6	43,6	19,6
7	12,5	48,3	28,2
8	16,5	44,7	22,9
9	16,0	35,7	23,3
10	16,1	49,3	23,1

Вариант 8

№	Уровень расходов	Собственная продукция	Трудоемкость
1	19,5	43,6	19,5
2	12,4	48,3	28,3
3	17,2	44,7	22,9
4	19,5	35,7	23,3
5	12,6	49,2	23,4
6	16,5	41,3	25,4
7	16,1	43,8	21,1
8	16,0	48,5	24,2
9	16,2	42,3	23,8
10	18,0	43,0	24,6

Вариант 9

№	Уровень расходов	Собственная продукция	Трудоемкость
1	11,7	44,2	28,5
2	18,3	42,4	18,1
3	18,2	38,9	18,8
4	15,6	36,4	22,5
5	17,4	35,2	19,8
6	13,8	45,5	24,8
7	15,0	35,2	22,5
8	18,6	41,6	18,5
9	16,2	42,2	21,5
10	15,7	40,4	22,5

Вариант 10

№	Уровень расходов	Собственная продукция	Трудоемкость
1	18,4	38,7	18,9
2	15,6	36,5	22,5
3	17,5	35,2	19,8
4	13,8	45,5	24,8
5	15,0	35,2	22,5
6	18,7	41,6	18,5
7	16,2	42,2	21,5
8	15,7	40,4	22,5
9	17,9	47,3	19,8
10	15,4	47,1	22,0

Вариант 11

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	13,8	45,5	24,6
2	15,0	35,2	22,5
3	18,6	41,6	18,5
4	16,2	42,3	21,5
5	15,7	40,4	22,5
6	17,9	47,4	19,0
7	15,3	47,1	22,0
8	16,3	43,2	20,9
9	17,7	39,1	18,7
10	16,8	38,2	20,0

Вариант 12

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	18,8	42,1	18,6
2	16,4	42,3	21,5
3	16,1	40,4	22,5
4	17,8	47,4	19,0
5	15,5	47,1	22,0
6	16,3	43,2	21,0
7	17,8	39,0	18,6
8	16,8	38,2	20,0
9	17,5	37,3	24,8
10	16,7	36,7	20,2

Вариант 13

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	13,8	48,0	24,5
2	14,8	46,4	23,0
3	16,9	42,3	20,0
4	16,8	39,4	20,5
5	14,8	42,3	22,3
6	18,0	40,1	18,9
7	17,5	39,4	20,0
8	15,7	39,1	22,1
9	15,3	43,2	20,1
10	14,9	44,7	23,1

Вариант 14

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	19,7	37,8	17,4
2	14,0	46,4	22,2
3	17,2	48,2	18,8
4	18,3	49,6	18,2
5	17,4	46,4	19,0
6	16,1	42,6	19,8
7	18,8	49,4	17,7
8	17,9	40,1	18,9
9	17,6	39,4	20,0
10	15,7	39,1	22,1

Вариант 15

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	16,8	40,4	18,8
2	17,0	42,3	20,0
3	14,8	40,1	22,1
4	17,9	39,4	20,1
5	16,9	39,1	23,1
6	19,7	43,2	21,6
7	14,0	44,5	17,4
8	17,1	45,7	22,5
9	18,2	37,8	18,7
10	17,4	46,4	18,2

Вариант 16

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	32,1	48,0	21,2
2	31,0	42,1	22,0
3	32,4	42,3	21,1
4	33,4	43,7	20,8
5	31,2	42,8	22,1
6	34,8	41,8	18,8
7	35,4	30,8	19,1
8	33,0	44,4	20,0
9	34,8	51.2	19,0
10	33,3	54,6	19,9

Вариант 17

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	35,4	30,0	19,0
2	33,0	44,4	20,0
3	34,8	51,2	19,1
4	33,3	51,6	19,9
5	36,1	57,6	15,4
6	38,3	53,2	17,4
7	30,6	57,6	22,3
8	32,1	58,3	21,4
9	37,6	55,7	18,4
10	34,8	55,7	20,1

Вариант 18

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	30,6	57,6	22,3
2	32,1	58,3	21,4
3	37,6	55,7	18,4
4	35,8	55,7	20,1
5	34,2	56,4	20,4
6	34,4	60,4	20,2
7	32,5	34,4	19,8
8	33,4	33,9	19,5
9	37,8	34,2	17,0
10	35,8	38,0	21,0

Вариант 19

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	34,2	30,3	17,5
2	37,2	31,1	18,6
3	38,2	28,5	18,2
4	29,4	36,7	18,9
5	37,2	38,7	20,1
6	34,5	45,7	17,5
7	35,0	41,5	16.0
8	43,7	39,0	16,9
9	31,9	42,7	17,8
10	37,3	39,8	24,8

Вариант 20

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	40,2	66,5	19,5
2	39,4	63,0	16,2
3	43,7	65,5	22,4
4	38,4	64,2	18,8
5	38,8	61,6	16,8
6	39,9	63,2	16,2
7	30,1	70,5	21,5
8	31,7	75,5	20,3
9	37,2	71,0	19,5
10	39,7	67,2	16,3

Вариант 21

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	31,51	36,6	28
2	30,20	51,6	27
3	29,10	48,3	22
4	32,79	63,2	38
5	26,44	74,6	30
6	37,16	60,8	35
7	26,04	32,2	24
8	31,91	53,2	25
9	27,13	41,6	59
10	29,30	76,1	43

Вариант 22

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	26,04	32,2	25
2	21,31	54,6	19
3	24,02	43,0	20
4	23,30	48,8	22
5	20,33	60,8	18
6	24,98	76,1	23
7	26,14	33,2	24
8	22,41	55,6	19
9	25,02	44,0	21
10	24,31	49,8	22

Вариант 23

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	17,9	82,1	9
2	23,8	76,2	10
3	19,7	80,3	12
4	15,1	84,9	14
5	18,8	81,2	8
6	18,5	81,5	9
7	14,8	85,2	14
8	22,2	77,8	10
9	17,8	82,2	12
10	17,7	80,1	11

Вариант 24

№	Уровень расходов	Собственная продукция	Трудоем-кость
1	27,3	38,2	35
2	21,3	54,6	19
3	24,0	43,0	27
4	22,9	48,7	28
5	21,3	59,3	21
6	25,5	75,0	22
7	28,3	39,2	26
8	22,3	55,6	20
9	25,0	44,0	28
10	23,9	49,7	29

Практическая работа №6
Исследование остатков эконометрической модели на автокорреляцию

6.1 Реализация эконометрических задач в оболочке электронных таблиц MS Excel

Данные для исследования остатков модели на автокорреляцию приведены в таблице 6.1.

Таблица 6.1 – Исходные данные для исследования

№	Y	X1	X2
1	4,17	9,74	17,01
2	4,55	5,56	33,85
3	3,27	10,95	25,14
4	3,35	4,14	31,95
5	4,68	10,04	32,45
6	3,35	9,16	24,62
7	3,66	4,91	24,78
8	3,06	3,91	30,88
9	2,74	6,01	20,7

Используя инструмент **Анализ данных**, построим регрессию (рисунок 6.1) и график остатков (рисунок 6.2).

Из графика видно, что зависимость остатков не наблюдается, соответственно можно сделать предварительный вывод об отсутствии автокорреляции остатков. Проверим наличие автокорреляции при помощи теста Дарбина – Уотсона. Сформулируем гипотезы.

H_0: С уровнем надежности $\alpha = 0{,}05$ можно утверждать, что автокорреляции не существует.

H_1: Автокорреляция существует.

Тест Дарбина – Уотсона основан на следующей статистике:

$$DW = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}.$$

Скопируем остатки на лист расчетов в MS Excel, преобразуем данные для расчетов (рисунки 6.3 и 6.4) и рассчитаем статистику *DW*.

ВЫВОД ИТОГОВ

Регрессионная статистика	
Множественнь	0,56121404
R-квадрат	0,31496119
Нормированнь	0,08661493
Стандартная о	0,64483601
Наблюдения	9

Дисперсионный анализ

	df	SS	MS	F	ачимость F
Регрессия	2	1,147074672	0,573537336	1,379314	0,321474
Остаток	6	2,494880883	0,415813481		
Итого	8	3,641955556			

	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
Y-пересечение	1,2125369	1,50673134	0,804746586	0,451667	-2,4743	4,899376	-2,4743	4,899376
Переменная X	0,12097943	0,088299981	1,370095814	0,219711	-0,09508	0,337042	-0,09508	0,337042
Переменная X	0,05851219	0,042338029	1,382024447	0,216214	-0,04509	0,16211	-0,04509	0,16211

ВЫВОД ОСТАТКА

Наблюдение	Предсказанное Y	Остатки
1	3,38616896	0,783831036
2	3,86582023	0,684179769
3	4,00825819	-0,738258195
4	3,58285627	-0,232856271
5	4,32589103	0,354108972
6	3,76127867	-0,411278668
7	3,25647802	0,403521976
8	3,49242296	-0,432422957
9	3,15082566	-0,410825661

Лист4 Лист5 Лист2 Лист3

Рисунок 6.1 – Построение регрессии

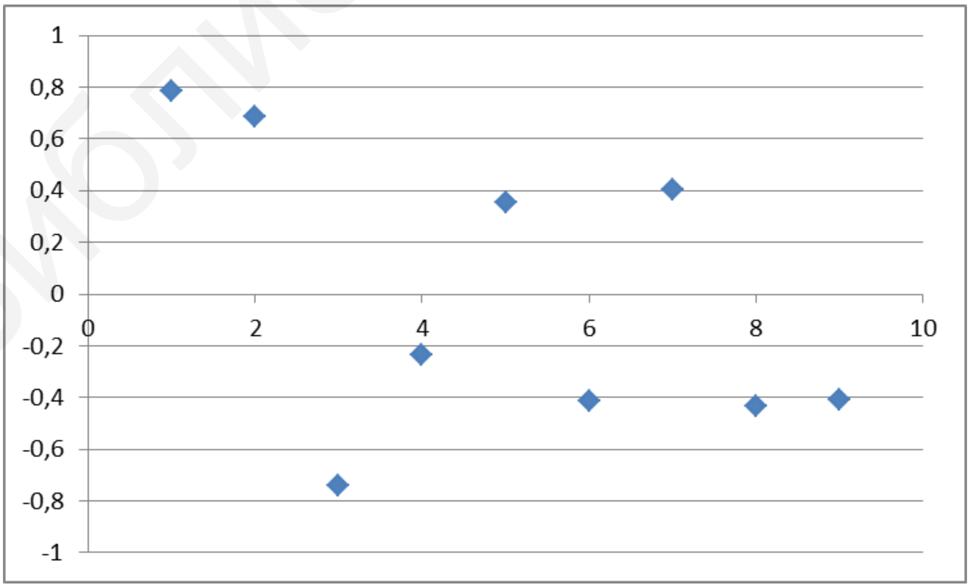


Рисунок 6.2 – График остатков

	A	B	C	D	E	F	G
1		Y	X1	X2	e	e^2	$(e_t$-$e_{t-1})$^2
2	1	4,17	9,74	17,01	0,783831	0,614391	
3	2	4,55	5,56	33,85	0,68418	0,468102	0,00993
4	3	3,27	10,95	25,14	-0,73826	0,545025	2,02333
5	4	3,35	4,14	31,95	-0,23286	0,054222	0,255431
6	5	4,68	10,04	32,45	0,354109	0,125393	0,344528
7	6	3,35	9,16	24,62	-0,41128	0,16915	0,585818
8	7	3,66	4,91	24,78	0,403522	0,16283	0,6639
9	8	3,06	3,91	30,88	-0,43242	0,18699	0,698804
10	9	2,74	6,01	20,7	-0,41083	0,168778	0,000466
11	Сумм					2,494881	4,582208
12	DW					1,839	

Рисунок 6.3 – Расчет статистики *DW* (значения)

	A	B	C	D	E	F	G
1		Y	X1	X2	e	e^2	$(e_t$-$e_{t-1})$^2
2	1	4,17	9,74	17,01	=Лист5!C26	=E2^2	
3	2	4,55	5,56	33,85	=Лист5!C27	=E3^2	=(E3-E2)^2
4	3	3,27	10,95	25,14	=Лист5!C28	=E4^2	=(E4-E3)^2
5	4	3,35	4,14	31,95	=Лист5!C29	=E5^2	=(E5-E4)^2
6	5	4,68	10,04	32,45	=Лист5!C30	=E6^2	=(E6-E5)^2
7	6	3,35	9,16	24,62	=Лист5!C31	=E7^2	=(E7-E6)^2
8	7	3,66	4,91	24,78	=Лист5!C32	=E8^2	=(E8-E7)^2
9	8	3,06	3,91	30,88	=Лист5!C33	=E9^2	=(E9-E8)^2
10	9	2,74	6,01	20,7	=Лист5!C34	=E10^2	=(E10-E9)^2
11	Сумм					=СУММ(F2:F10)	=СУММ(G2:G10)
12	DW					=G11/F11	

Рисунок 6.4 – Расчет статистики *DW* (формулы)

Для обоснованного вывода об автокорреляции на заданном уровне значимости α используют пороговые значения d_l, d_u, вычисленные для некоторых фиксированных значений числа наблюдений n и количества экзогенных переменных в исходной модели. В соответствии с тестом Дарбина – Уотсона принимаются следующие решения (таблица 6.2).

Таблица 6.2 – Принятие решений по статистике Дарбина – Уотсона

Значение статистики DW	Вывод
$4-d_l<DW<4$	Гипотеза H_0 отвергается, есть отрицательная корреляция.
$4-d_u<DW<4-d_l$	Неопределенность.
$d_u<DW<4-d_u$	Гипотеза H_0 не отвергается.
$d_l<DW<d_u$	Неопределенность.
$0<DW<d_l$	Гипотеза H_0 отвергается, есть положительная корреляция

Для нашего примера $d_l=0,629$, $d_u=1,699$. Имеем следующие данные для принятия решения (рисунок 6.5)

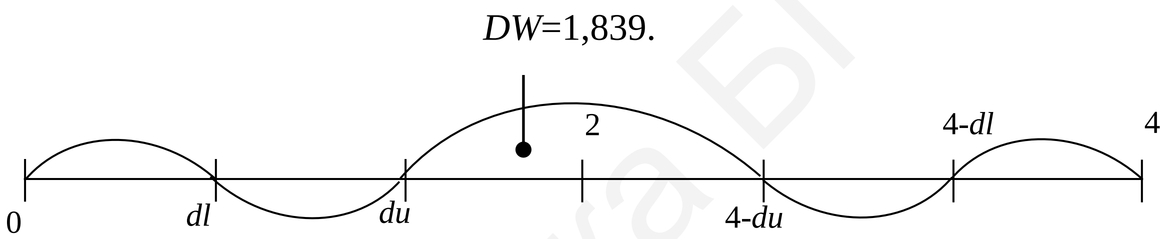


Рисунок 6.5 – *DW*-статистика

Статистика Дарбина – Уотсона попала в зону положительной автокорреляции, соответственно гипотеза H_0 отвергается, существует положительная автокорреляция.

6.2 Варианты заданий

На основе имеющихся данных (из подраздела 3.13) составить уравнение множественной регрессии, рассчитать теоретические данные результирующего показателя и остатки модели. Проверить остатки на автокорреляцию графическим методом и при помощи теста Дарбина – Уотсона.

Практическая работа №7
Косвенный метод наименьших квадратов (КМНК)

7.1 Понятие косвенного метода наименьших квадратов

Система регрессий называется полной, если:

- она имеет столько регрессий, сколько в ней эндогенных значений;
- она имеет все переменные, которые имеют существенное влияние на взаимозависимые эндогенные значения;
- систему можно решить относительно эндогенных переменных.

Для оценки параметров такой модели используется косвенный метод наименьших квадратов (КМНК), если она идентифицирована.

Рассмотрим суть КМНК для полной системы двух регрессий. Допустим, что между эндогенными и экзогенными переменными существует линейная зависимость:

$$\begin{cases} Y_1 = b_{12}Y_2 + a_{10} + a_{11}X_1 + e_1, \\ Y_2 = b_{21}Y_1 + a_{20} + a_{22}X_2 + e_2, \end{cases} \tag{7.1}$$

где Y_1, Y_2 – векторы эндогенных переменных;

X_1, X_2 – векторы экзогенных переменных;

e_1, e_2 – векторы отклонений (остатков).

Перепишем модель в виде

$$\begin{cases} -Y_1 + b_{12}Y_2 + a_{10} + a_{11}X_1 + e_1 = 0, \\ -Y_2 + b_{21}Y_1 + a_{20} + a_{22}X_2 + e_2 = 0 \end{cases} \tag{7.2}$$

или в матричном виде

$$Y \cdot B + X \cdot A + e = 0, \tag{7.3}$$

где $B = \begin{pmatrix} -1 & b_{21} \\ b_{12} & 1 \end{pmatrix}$ $A = \begin{pmatrix} a_{10} & a_{20} \\ a_{11} & 0 \\ 0 & a_{22} \end{pmatrix}$ $Y = \begin{pmatrix} y_{11} & y_{21} \\ y_{12} & y_{22} \\ \vdots & \vdots \\ y_{1n} & y_{2n} \end{pmatrix}$

$$X = \begin{pmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} \end{pmatrix} e = \begin{pmatrix} e_{11} & e_{21} \\ e_{12} & e_{22} \\ \vdots & \vdots \\ e_{1n} & e_{2n} \end{pmatrix}.$$

Модель (7.2) или (7.3) называется структурной формой эконометрической модели. Если определитель $\det[B] \neq 0$, то систему регрессий можно решить относительно эндогенных переменных. Перемножим модель (7.3) на матрицу B^{-1} справа:

$$Y = -X \cdot A \cdot B^{-1} - e \cdot B^{-1}.$$

Если обозначить

$$C = A \cdot B^{-1}, \qquad (7.4)$$
$$U = -e \cdot B^{-1},$$

то получим модель, которая называется прогнозной (сводной) формой эконометрической модели

$$Y = X \cdot C + U, \qquad (7.5)$$

где $C = \begin{pmatrix} c_{10} & c_{20} \\ c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix}$ – матрица коэффициентов прогнозной формы модели, или в развернутом виде

$$\begin{cases} Y_1 = c_{10} + c_{11}X_1 + c_{12}X_2 + U_1, \\ Y_2 = c_{20} + c_{21}X_1 + c_{22}X_2 + U_2. \end{cases} \qquad (7.6)$$

Элементы матриц A и B называют структурными коэффициентами, а матрицы C – коэффициентами прогнозной формы.

Оценкой прогнозной модели является модель

$$\hat{Y} = X \cdot \hat{C}. \qquad (7.7)$$

Оценки прогнозных коэффициентов получают методом 1МНК, т. е. по формуле

$$\hat{C} = \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y, \qquad (7.8)$$

где $\hat{C}$ – матрица оценок параметров прогнозной модели.

Для оценки структурных коэффициентов используем матричное уравнение, где

$$\hat{A} = -\hat{C} \cdot \hat{B}, \qquad (7.9)$$

которое получили из уравнения (7.9), или в развернутом виде (для упрощения записи знак ^ опустим)

$$\begin{pmatrix} a_{10} & a_{20} \\ a_{11} & 0 \\ 0 & a_{22} \end{pmatrix} = \begin{pmatrix} c_{10} & c_{20} \\ c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix} \cdot \begin{pmatrix} -1 & b_{21} \\ b_{12} & 1 \end{pmatrix}. \qquad (7.10)$$

Расписав произведение матриц, получим шесть уравнений:

$$
\begin{cases}
-c_{10} + c_{20}b_{12} = -a_{10}, \\
-c_{11} + c_{21}b_{12} = -a_{11}, \\
-c_{12} + c_{22}b_{12} = 0, \\
c_{10}b_{21} - c_{20} = -a_{20}, \\
c_{11}b_{21} - c_{21} = 0, \\
c_{12}b_{21} - c_{22} = -a_{22}.
\end{cases}
\tag{7.11}
$$

Если система (7.11) имеет единственное решение, то параметры структурной формы регрессий выражаются через параметры прогнозной формы регрессии однозначно по формулам:

$$
\begin{aligned}
b_{12} &= \frac{c_{12}}{c_{22}}, & b_{21} &= \frac{c_{21}}{c_{11}}, \\
a_{11} &= \frac{c_{11}c_{22} - c_{21}c_{12}}{c_{22}}, & a_{22} &= \frac{c_{11}c_{22} - c_{21}c_{12}}{c_{11}}, \\
a_{10} &= \frac{c_{10}c_{22} - c_{12}c_{20}}{c_{22}}, & a_{20} &= \frac{c_{11}c_{20} - c_{21}c_{10}}{c_{11}}.
\end{aligned}
\tag{7.12}
$$

Эконометрическая модель называется идентифицированной, если параметры структурной формы модели однозначно выражаются через параметры прогнозной модели.

Эконометрическая модель называется неидентифицируемой, если количество оцениваемых парметров структурной формы модели больше количества оцениваемых параметров модели, т. е. больше, чем количество уравнений.

7.2 Алгоритм построения и анализа эконометрической модели методом КМНК в среде табличного процессора MS Excel

7.2.1 Идентификация переменных и спецификация модели.

7.2.2 Проверка условий идентифицируемости для каждого уравнения структурной формы.

7.2.3 Составление прогнозной формы модели и оценка ее параметров 1МНК.

7.2.4 Определение оценок параметров структурной формы модели.

7.2.5 Оценка адекватности прогнозной формы эконометрической модели.

7.2.6 Прогнозирование по оцененной прогнозной модели.

7.2.7 Экономический анализ по оцененной структурной модели.

7.3 Реализация косвенного метода наименьших квадратов для решения СОУ в оболочке электронных таблиц MS Excel

На основе статистических данных за 10 лет про экспорт, импорт, национальный доход и средний оборот внешней торговли стран, с которыми поддерживаются внешнеэкономические отношения, оценить параметры прогнозной и структурной форм эконометрической модели «экспорт – импорт» блока внешней торговли *N*-й страны. Найти значения прогноза экспорта – импорта, если прогнозные значения национального дохода и среднего оборота внешней торговли равны соответственно 115 и 11 усл. ден. ед.

Таблица 7.1 – Исходные данные и оцененные значения эндогенных переменных

	A	B	C	D	E	F	G	H	I	J	K	L
1												
2		№	X_1	X_2	Y_1	Y_2	$\hat{Y}_1$	$\hat{Y}_2$	$Y_1-\hat{Y}_1$	$Y_2-\hat{Y}_2$	Y_1-$Y_{ср}$	Y_2-$Y_{ср}$
3		1	70	5,5	55	73	52,6	75,6	2,4	-2,6	-45,0	-38,0
4		2	75	6,1	63	86	64,3	84,1	-1,3	1,9	-37,0	-25,0
5		3	77	6,9	71	94	72,1	95,7	-1,1	-1,7	-29,0	-17,0
6		4	80	7,4	80	103	79,9	102,8	0,1	0,2	-20,0	-8,0
7		5	85	7,9	93	111	91,1	109,9	1,9	1,1	-7,0	0,0
8		6	92	8,3	103	117	105,0	115,3	-2,0	1,7	3,0	6,0
9		7	99	8,8	117	122	119,5	122,2	-2,5	-0,2	17,0	11,0
10		8	103	9,1	128	130	127,9	126,3	0,1	3,7	28,0	19,0
11		9	107	9,8	139	135	138,5	136,3	0,5	-1,3	39,0	24,0
12		10	112	10,2	151	139	149,1	141,9	1,9	-2,9	51,0	28,0
13		Сумм	900	80	1000	1110	1000	1110				
14		Суммкв							26,8	41,3	9888,0	4300,0
15		Срзнач			100	111						
16		Прогноз	115	11			156,205	153,26				

Поставленную задачу решим в среде Excel по указанному выше алгоритму.

7.3.1 Идентификация переменных и спецификация модели. Пусть векторы эндогенных переменных:

Y_1 – экспорт;

Y_2 – импорт;

и экзогенных переменных:

X_1 – национальный доход страны;

X_2 – внешний товарооборот стран, с которыми поддерживаются внешнеэкономические отношения.

Допустим, что система зависимостей между этими макроэкономическими показателями имеет вид:

- структурная форма:

$$\begin{cases} Y_1 = b_{12}Y_2 + a_{10} + a_{11}X_1 + e_1, \\ Y_2 = b_{21}Y_1 + a_{20} + a_{22}X_2 + e_2; \end{cases}$$

- прогнозная форма:

$$\begin{cases} Y_1 = c_{10} + c_{11}X_1 + c_{12}X_2 + U_1, \\ Y_2 = c_{20} + c_{21}X_1 + c_{22}X_2 + U_2. \end{cases}$$

7.3.2 Проверка условий идентифицируемости для каждого уравнения структурной формы. Для проверки выполнения условий идентифицруемости модели составим для каждой регрессии соотношения (7.4):

- для первой регрессии: 2 – 1 = 2 – 1;
- для второй регрессии: 2 – 1 = 2 – 1.

Условие индентифцируемости регрессий выполнено, поэтому оценки параметров структурной формы модели можно определить косвенным методом наименьших квадратов (КМНК).

	B	C	D	E	F	G	H	I	J	K	L	M
17												
18		1	70	5,5		55	73		10	900	80	
19		1	75	6,1		63	86	$X^T*X=$	900	82946	7400,4	
20		1	77	6,9		71	94		80	7400,4	661,46	
21		1	80	7,4		80	103					
22	X=	1	85	7,9	Y=	93	111		6,217427	-0,205	1,541537	
23		1	92	8,3		103	117	$(X^T*X)^{-1}=$	-0,205	0,013404	-0,12517	
24		1	99	8,8		117	122		1,541537	-0,12517	1,21549	
25		1	103	9,1		128	130					
26		1	107	9,8		139	135		1000	1110		
27		1	112	10,2		151	139	$X^T*Y=$	94375	102718		
28									8455,1	9182,3		
29												
30		1	1	1	1	1	1	1	1	1	1	
31	$X^T=$	70	75	77	80	85	92	99	103	107	112	
32		5,5	6,1	6,9	7,4	7,9	8,3	8,8	9,1	9,8	10,2	
33												
34		-95,3081	-0,6747					Прогнозная форма модели				
35	$C^T=$	1,677364	-0,06661		$Y_1=$	-95,3081	+	1,677364	$*X_1+$	5,543161	$*X_2$	
36		5,543161	14,70868		$Y_2=$	-0,6747	+	-0,06661	$*X_1+$	14,70868	$*X_2$	
37												

Рисунок 7.1 – Оценка параметров прогнозной формы 1МНК

7.3.3 Составление прогнозной формы модели и оценка ее параметров 1МНК. Прогнозная матричная форма модели имеет вид (7.10), т. е. $Y = X \cdot C + U$, оценку параметров этой модели получим методом 1МНК, т. е. по формуле (7.13): $\hat{C} = \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y$.

Исходя из формулы (7.6), составим оцененную прогнозную модель в диапазоне F35:L36.

Таким образом, оцененная прогнозная модель имеет вид

$$\begin{cases} Y_1 = -95{,}308 + 1{,}677 X_1 + 5{,}543 X_2, \\ Y_2 = -0{,}675 Y_1 - 0{,}067 X_1 + 14{,}709 X_2. \end{cases}$$

7.3.4 Определение оценок параметров структурной формы модели. Оценки параметров структурной формы модели вычислим по формулам (7.12) в ячейках C39, C40, E39, E40, G39, G40 (рисунок 7.2).

	B	C	D	E	F	G	H	I
38								
39		b_{12}=	0,377	a_{10}=	-95,054	a_{11}=	1,702	
40		b_{21}=	-0,040	a_{20}=	-4,459	a_{22}=	14,929	
41								
42		-1	0,377		-95,3081	1,677364	5,543161	
43	B=	-0,040	-1	C=	-0,6747	-0,06661	14,70868	
44								
45		-95,054	1,702	0		95,05379	-1,70247	0
46	A=	-4,459	0	14,929	B*C=	4,4594	0	-14,9288
47					Структурная форма модели			
48	Y_1=	0,377	$*Y_2+$	-95,054	+	1,702	$*X_2$	
49	Y_2=	-0,040	$*Y_1+$	-4,459	+	14,929	$*X_2$	
50				Оценка адекватности модели				
51	I регрессия		R^2_1=	0,997	$F_{1расч}$=	1285,943	Модель адекватна	
52	II регрессия		R^2_2=	0,990	$F_{2расч}$=	361,1912	Модель адекватна	
53					$F_{табл}$=	**4,737414**		
54								

Рисунок 7.2 – Определение оценок параметров структурной формы модели. Оценка адекватности модели

Оцененную структурную модель составим, исходя из формулы (7.5). Таким образом, оцененная структурная модель имеет вид

$$\begin{cases} \hat{Y}_1 = 0{,}377Y_2 - 95{,}054 + 1{,}702X_1, \\ \hat{Y}_2 = -0{,}040Y_1 - 4{,}459 + 14{,}929X_2. \end{cases}$$

7.3.5 Оценка адекватности прогнозной формы эконометрической модели:

1 Оценка адекватности модели по коэффициенту множественной детерминации.

Для предварительной оценки адекватности полученной прогнозной модели вычислим коэффициенты множественной детерминации для первой и второй регрессии по формуле

$$R^2 = 1 - \frac{\sum_{i=1}^{n} e_i^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}.$$

По алгоритму вычисления коэффициента множественной детерминации и корреляции, который был приведен в практической работе №3, произведем вычисления в следующем порядке:

- используя прогнозную форму модели в диапазоне G3:H12 (рисунок 7.1) вычислим расчетные значения эндогенных переменных;

- отклонения расчетных значений от статистических эндогенных Y_1 и Y_2 значений рассчитываются в диапазоне I3:J12;

- отклонения статистических данных от средних эндогенных значений $\bar{Y}_1$ и $\bar{Y}_2$ рассчитываются в диапазоне K3:L12.

Анализ полученных данных приводит к таким выводам:

- поскольку коэффициент множественной детерминации для первой регрессии равен 0,9973, то это свидетельствует о том, что вариация экспорта на 99,37 % определяется вариацией импорта и национального дохода страны и только на 0,27 % – вариацией факторов, которые не учтены в регрессии;

- поскольку коэффициент множественной детерминации для второй равен 0,9904, то это свидетельствует о том, что вариация импорта на 99,04 % определяется вариацией экспорта и внешнего товарооборота страны и на 0,96 % определяется вариацией факторов, которые не учтены в регрессии.

Таким образом, полученную эконометрическую модель можно считать адекватной статистическим данным.

2 Оценка адекватности модели по критерию Фишера.

Проверку адекватности модели по критерию Фишера проведем относительно алгоритма, приведенного в практической работе №2. Для числа степеней свободы $k_1 = 2$ и $k_2 = 7$ и доверительной вероятности P = 0,95 критическое значение F-критерия равно 4,74. Расчетное значение F-критерия значительно больше критического, поэтому с надежностью P = 0,95 можно считать, что принятая модель «экспорт – импорт» адекватна статистическим данным и на основе этой модели можно проводить экономический анализ и прогнозирование.

7.3.6 Прогнозирование по оцененной прогнозной модели. Для прогнозирования точечных оценок прогноза используется система регрессий в прогнозной форме. По условию примера прогнозные значения экзогенных значений равны $x_{1p} = 115$, $x_{2p} = 11$. Введем эти значения соответственно в ячейки C16:D16 (таблица 7.1) и копированием формул в ячейках G16:H16 получим прогнозные значения эндогенных переменных, т. е. прогнозное значение экспорта 158,6 усл. ден. ед. и импорта 153,5 усл. ден. ед.

7.3.7 Экономический анализ по оцененной структурной модели. Анализ структурной формы модели приводит к следующим результатам:

- при увеличении Y_2 (импорта) на единицу при неизменных других значениях Y_1 (экспорт) увеличится на 0,38 усл. ден. ед.;

- при увеличении X_1 (национального дохода) на единицу при неизменных значениях Y_1 (экспорт) увеличится на 1,70 усл. ден. ед.;

- при увеличении Y_1 (экспорт) на единицу при неизменных других значениях Y_2 (импорт) уменьшится на 0,04 усл. ден. ед.;

- при увеличении X_2 (среднего внешнего товарооборота стран блока) на единицу при неизменных других значениях Y_2 (импорта) увеличится на 14,93 усл. ден. ед.

7.4 Варианты заданий

Итоговыми показателями эффективности работы строительных организаций является общая реализация строительной продукции (объем сданных строительных работ) и общая рентабельность. На основе статистических данных по 12 строительно-монтажным организациям об общей реализации строительной продукции (Y_1), общей рентабельности (Y_2), доле сборных конструкций, деталей и изделий в себестоимости строительно-монтажных работ (X_1), выработке на одного рабочего (X_2) оценить параметры прогнозной и структурной форм эконометрической модели, допуская, что структурная форма модели имеет вид

$$\begin{cases} Y_1 = b_{12}Y_2 + a_{10} + a_{11}X_1 + e_1, \\ Y_2 = b_{21}Y_1 + a_{20} + a_{22}X_2 + e_2. \end{cases}$$

Руководство объединения хотело бы знать, насколько изменится общая реализация строительной продукции и общая рентабельность в СМУ№5, если часть сборных конструкций, деталей и изделий в себестоимости строительно-монтажных работ и выработка на одного рабочего увеличатся на 5 %.

Вариант	**Y_1 – общая реализация строительной продукции (млн усл. ден. ед.)**											
	1	2	3	4	5	6	7	8	9	10	11	12
1	82,7	85,4	87,7	92,3	98,5	105,4	106,5	108,1	115,6	124,3	127,7	132,6
2	84,1	86,8	89,1	93,7	99,9	106,8	107,9	109,5	117	125,7	129,1	134
3	85	87,7	90	94,6	100,8	107,7	108,8	110,4	117,9	126,6	130	134,9
4	89,8	92,5	94,8	99,4	105,6	112,5	113,6	115,2	122,7	131,4	134,8	139,7
5	86,3	89	91,3	95,9	102,1	109	110,1	111,7	119,2	127,9	131,3	136,2
6	89,2	91,9	94,2	98,8	105	111,9	113	114,6	122,1	130,8	134,2	139,1
7	84,5	87,2	89,5	94,1	100,3	107,2	108,3	109,9	117,4	126,1	129,5	134,4
8	86,2	88,9	91,2	95,8	102	108,9	110	111,6	119,1	127,8	131,2	136,1
9	85,3	88	90,3	94,9	101,1	108	109,1	110,7	118,2	126,9	130,3	135,2
10	86,7	89,4	91,7	96,3	102,5	109,4	110,5	112,1	119,6	128,3	131,7	136,6
11	87,6	90,3	92,6	97,2	103,4	110,3	111,4	113	120,5	129,2	132,6	137,2
12	92,4	95,1	97,4	102	108,2	115,1	116,2	117,8	125,3	134	137,4	142,3
13	87,1	89,8	92,1	96,7	102,9	109,8	110,9	112,5	120	128,7	132,1	137
14	88,8	91,5	93,8	98,4	104,6	111,5	112,6	114,2	121,7	130,4	133,8	138,7
15	87,9	90,6	92,9	97,5	103,7	110,6	11,7	113,3	120,8	129,5	132,9	137,8
16	89,3	92	94,3	98,9	105,1	112	113,1	114,7	122,2	130,9	134,3	139,2
17	90,2	92,9	95,2	99,8	106	112,9	114	115,6	123,1	131,8	135,2	140,1
18	95,9	97,7	100	104,6	110,8	117,7	118,8	120,4	127,9	136,6	140	144,9
19	92,4	95,1	97,4	102	108,2	115,1	116,1	117,8	125,3	134	137,4	142,3
20	95,3	98	100,3	104,9	111,1	118	119,1	120,7	128,2	136,9	140,3	145,2
21	90,6	93,3	95,6	100,2	106,4	113,3	114,4	116	123,5	132,2	135,5	140,5
22	92,3	95	97,3	101,9	108	115,1	116,1	117,7	125,2	133,9	137,3	142,2
23	87,1	89,8	92,1	96,7	102,9	109,8	110,9	112,5	120	128,7	132,1	137
24	88,8	91,5	93,8	98,4	104,6	11,5	112,6	114,2	121,7	130,4	133,8	138,7

Вариант	Y_2 – общая рентабельность (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
1	7,43	8,84	9,8	8,54	9,56	9,15	8,31	9,21	10,4	9,98	10,6	10,9
2	4,42	5,12	5,6	4,97	5,48	5,78	5,86	6,31	5,93	6,19	6,54	6,83
3	4,62	5,32	5,8	5,17	5,68	5,98	6,06	6,51	6,13	6,39	6,74	7,03
4	7,27	7,97	8,45	7,82	8,33	8,63	8,71	9,16	8,78	9,04	9,39	9,68
5	5,52	6,22	6,7	6,07	6,58	6,88	6,96	7,41	7,03	7,29	7,64	7,93
6	6,97	7,67	8,15	7,52	8,03	8,33	8,41	8,86	8,48	8,74	9,09	9,38
7	7,67	8,37	8,85	8,22	8,73	9,03	9,11	9,56	9,18	9,44	9,79	10,08
8	8,12	8,82	9,3	8,67	9,18	8,48	9,56	10,01	9,63	9,89	10,24	10,53
9	5,02	5,72	6,2	5,57	6,08	6,38	6,46	6,91	6,53	6,79	7,14	7,43
10	5,72	6,42	6,9	6,27	6,78	7,08	7,16	7,61	7,23	7,49	7,84	8,13
11	6,17	6,87	7,35	6,72	7,23	7,53	7,61	8,06	7,68	7,94	8,29	8,58
12	4,93	5,29	5,53	5,21	5,47	5,61	5,65	5,88	5,69	5,82	5,99	6,14
13	4,06	4,41	4,65	4,34	4,59	4,74	7,78	5	4,81	4,95	5,12	5,27
14	7,11	7,81	8,29	7,66	8,17	8,47	8,55	9	8,62	8,88	9,23	9,52
15	1,96	2,31	2,55	2,24	2,49	2,64	2,68	2,9	2,71	2,85	3,02	3,17
16	4,52	5,22	5,7	5,07	5,58	5,88	5,96	6,41	6,03	6,29	6,64	6,93
17	7,62	8,32	8,8	8,17	8,68	8,98	9,06	9,51	9,13	9,39	9,74	10,03
18	6,21	6,56	6,8	6,49	6,74	6,89	6,93	7,15	6,96	7,1	7,27	7,42
19	8,77	9,47	9,95	9,32	9,83	10,13	10,21	10,66	10,28	10,54	10,89	11,18
20	7,53	7,89	8,13	7,81	8,07	8,21	8,25	8,48	8,29	8,42	8,59	8,74
21	4,56	4,91	5,15	4,84	5,09	5,24	5,28	5,5	5,31	5,45	5,62	5,77
22	5,3	5,66	5,9	5,58	5,84	5,98	6,02	6,25	6,06	6,19	6,36	6,54
23	2,73	2,91	3,03	2,87	3	3,07	3,09	3,2	3,11	3,17	3,26	3,33
24	4,01	4,36	4,6	4,29	4,54	4,69	4,73	4,95	4,76	4,9	5,07	5,22

Вариант	X_1 – доля сборных конструкций, деталей и изделий в себестоимости строительно-монтажных работ (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
1	30,22	30,74	31,05	31,09	31,21	31,55	31,64	31,75	32,09	32,22	32,6	32,88
2	30,26	30,78	31,09	31,14	31,26	31,59	31,69	31,79	32,13	32,26	32,65	32,93
3	30,5	31,02	31,33	31,5	31,83	31,93	32,03	32,37	32,5	32,89	32,89	33,17
4	30,33	30,85	31,16	31,2	31,32	31,66	31,75	31,86	32,2	32,33	32,71	32,99
5	30,47	30,99	31,3	31,35	31,47	31,8	31,9	32	32,34	32,47	32,86	33,14
6	30,24	30,76	31,07	31,11	31,23	31,57	31,66	31,77	32,11	32,24	32,62	32,9
7	30,32	30,84	31,15	31,2	31,32	31,65	31,75	31,85	32,19	32,32	32,71	32,98
8	30,28	30,8	31,11	31,15	31,27	31,61	31,7	31,81	32,15	32,28	32,66	32,94
9	32,82	33,34	33,65	33,69	33,81	34,15	34,24	34,35	34,69	34,82	35,2	35,48
10	32,82	33,34	33,65	33,69	33,81	34,15	34,24	34,35	34,69	34,82	35,2	35,48
11	33,1	33,62	33,93	33,98	34,1	34,43	34,53	34,63	34,97	35,1	35,49	35,76
12	32,93	33,45	33,76	33,8	33,92	34,26	34,35	34,46	34,8	34,93	35,31	35,59
13	33,07	33,59	33,9	33,95	34,07	34,4	34,5	34,6	34,94	35,07	35,46	35,74

Вариант	X_1 – доля сборных конструкций, деталей и изделий в себестоимости строительно-монтажных работ (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
14	32,84	33,36	33,67	33,71	33,83	34,17	34,26	34,37	34,71	34,84	35,22	35,5
15	32,92	33,44	33,75	33,8	33,92	34,25	34,35	34,45	34,79	34,92	35,31	35,59
16	32,88	33,4	33,71	33,75	33,87	34,21	34,3	34,41	34,75	34,88	35,26	35,54
17	35,42	35,94	36,25	36,29	36,41	36,75	36,84	36,95	37,29	37,42	37,8	38,08
18	35,46	35,98	36,29	36,34	36,46	36,79	36,89	36,99	37,33	37,46	37,8	38,08
19	35,7	36,22	36,53	36,58	36,7	37,03	37,13	37,23	37,57	37,7	3809	38,37
20	36,43	36,95	37,26	37,3	37,42	37,76	37,85	37,96	38,3	38,43	38,81	39,09
21	36,57	37,09	37,4	37,45	37,57	37,9	38	37,1	38,44	38,57	38,96	39,24
22	36,34	36,86	37,17	37,21	37,33	37,67	37,76	37,87	38,21	38,34	38,72	39
23	36,42	36,94	37,25	37,3	37,42	37,75	37,85	37,95	38,29	38,42	38,81	39,09
24	36,6	36,9	37,2	37,3	37,4	37,6	37,8	37,9	38,2	38,5	38,9	39

Вариант	X_2 – выработка на одного рабочего (тыс. усл. ден. ед.)											
	1	2	3	4	5	6	7	8	9	10	11	12
1	8,71	8,86	8,93	8,99	9,09	9,16	9,26	9,31	9,4	9,44	9,59	9,78
2	10,11	10,26	10,33	10,39	10,49	10,56	10,66	10,71	10,8	10,84	10,99	11,18
3	11,01	11,16	11,23	11,29	11,39	11,46	11,56	11,61	11,7	11,74	11,89	12,08
4	15,81	15,96	16,03	16,09	16,19	16,26	16,36	16,41	16,5	16,54	16,69	16,88
5	12,31	12,46	12,53	12,59	12,69	12,76	12,86	12,91	13	13,04	13,19	13,38
6	15,21	15,36	15,43	15,49	15,59	15,66	15,76	15,81	15,9	15,94	16,09	16,28
7	10,51	10,66	10,73	10,79	10,89	10,96	11,06	11,11	11,2	11,24	11,39	11,58
8	12,21	12,36	12,43	12,49	12,59	12,66	12,73	12,81	12,9	12,94	13,09	13,28
9	11,31	11,46	11,53	11,59	11,69	11,76	11,86	11,91	12	12,04	12,19	12,39
10	12,71	12,86	12,93	12,99	13,09	13,16	13,26	13,31	13,4	13,44	13,59	13,78
11	13,61	13,76	13,83	13,89	13,99	14,06	14,16	14,21	14,3	14,34	14,49	14,68
12	18,41	18,56	18,63	18,69	18,79	18,86	18,96	19,01	19,1	19,14	19,29	19,48
13	14,91	16,06	15,13	15,19	15,29	15,36	15,46	15,51	15,6	15,64	15,79	15,98
14	17,81	17,96	18,03	18,09	18,19	18,26	18,36	15,41	18,5	18,54	18,69	18,88
15	13,11	13,26	13,33	13,39	13,49	13,56	13,66	13,71	13,8	13,84	13,99	14,18
16	14,81	14,96	15,03	15,09	15,19	15,26	15,36	15,41	15,5	15,54	15,69	15,88
17	13,91	14,06	14,13	14,19	14,29	14,36	14,46	14,51	14,6	14,64	14,79	14,98
18	15,31	15,46	15,53	15,59	15,69	15,76	15,86	15,91	16	16,04	15,19	15,38
19	16,21	16,36	16,43	16,49	16,59	16,66	16,76	16,81	16,8	16,94	17,09	17,28
20	21,01	21,16	21,23	21,29	21,39	21,46	21,56	21,61	21,7	21,74	21,89	22,08
21	18,41	18,56	18,63	18,69	18,79	18,86	18,96	19,01	19	19,14	19,29	19,48
22	21,31	21,46	21,53	21,59	21,69	21,76	21,86	21,91	22	22,04	22,19	22,38
23	16,61	16,76	16,83	16,89	16,99	17,06	17,16	17,21	17,2	17,34	17,49	17,68
24	18,31	18,46	18,53	18,59	18,69	18,76	18,86	18,91	18,9	19,04	19,19	19,38

Практическая работа №8
Двухшаговый метод наименьших квадратов (2МНК)

8.1 Понятие двухшагового метода наименьших квадратов

Если уравнения структурной формы сверхидентифицированы, то косвенный метод наименьших квадратов использовать нельзя, а пользоваться 1МНК нецелесообразно. Одним из методов, которые разработаны специально для таких моделей, является двухшаговый метод наименьших квадратов (2МНК).

Рассмотрим систему двух взаимозависимых линейных регрессий:

$$\begin{cases} Y_1 = b_{12}Y_2 + a_{10} + a_{11}X_1 + a_{12}X_2 + e_1, \\ Y_2 = b_{21}Y_1 + a_{20} + a_{12}X_1 + a_{22}X_2 + e_2, \end{cases} \tag{8.1}$$

где Y_1, Y_2 – векторы эндогенных переменных;

X_1, X_2 – векторы экзогенных переменных;

e_1, e_2 – векторы отклонений (остатков).

Проверим систему на идентифицируемость. Для первого уравнения условие запишется следующим образом:

$$4 - 2 > 2 - 1,$$

т. е. невозможно оценить параметры структурной формы КМНК.

Алгоритм 2МНК состоит из двух шагов:

- первый шаг состоит в оценке 1МНК параметров прогнозной формы модели;

- второй шаг заключается в оценке параметров структурной формы модели.

Алгоритм двухшагового МНК

Первый шаг 2МНК.

1 Проверить каждое уравнение структурной формы на идентифицируемость. Если все уравнения сверхидентифицируемы, то используется 2МНК.

2 Записать прогнозную форму модели в матричном виде, т. е. модель (7.5):

$$Y = X \cdot C + U,$$

где $C = \begin{pmatrix} c_{10} & c_{20} \\ c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix}$ – матрица коэффициентов прогнозной формы модели, допуская, что определитель матрицы $B = \begin{pmatrix} -1 & b_{21} \\ b_{12} & 1 \end{pmatrix}$ не равен нулю.

3 Оценить параметры прогнозной формы модели методом 1МНК по формуле (7.8):

$$\hat{C} = \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y,$$

если определитель $\det\left[X^T \cdot X\right] \neq 0$.

Второй шаг 2МНК.

4 Найти матрицу расчетных значений эндогенных переменных по формуле

$$\hat{Y} = X \cdot \hat{C} = \begin{pmatrix} \hat{y}_{11} & \hat{y}_{21} \\ \hat{y}_{12} & \hat{y}_{22} \\ \vdots & \vdots \\ \hat{y}_{1n} & \hat{y}_{2n} \end{pmatrix}.$$

5 Составить систему структурных регрессий, принимая значения Y_i, которые находятся справа в системе регрессий (8.1), предопределенными, т. е. заменить их на $\hat{Y}_i$:

$$\begin{cases} Y_1 = b_{12}\hat{Y}_2 + a_{10} + a_{11}X_1 + a_{12}X_2 + e_1, \\ Y_2 = b_{21}\hat{Y}_1 + a_{20} + a_{12}X_1 + a_{22}X_2 + e_2. \end{cases} \tag{8.2}$$

6 Оценить параметры каждой из регрессий (8.2) методом 1МНК. Векторы оценок параметров первой и второй регрессий можно получить по соответствующим формулам:

$$\hat{d}_1 = Z_1^{-1} \cdot \begin{pmatrix} X^T \cdot Y_1 \\ \hat{Y}_1^T \cdot Y_1 \end{pmatrix}, \qquad \hat{d}_2 = Z_2^{-1} \cdot \begin{pmatrix} X^T \cdot Y_2 \\ \hat{Y}_2^T \cdot Y_2 \end{pmatrix}, \tag{8.3}$$

где $\hat{d}_1 = \begin{pmatrix} \hat{a}_{10} \\ \hat{a}_{11} \\ \hat{a}_{12} \\ \hat{b}_{12} \end{pmatrix}, \hat{d}_2 = \begin{pmatrix} \hat{a}_{20} \\ \hat{a}_{21} \\ \hat{a}_{22} \\ \hat{b}_{22} \end{pmatrix}$ – векторы оценок параметров соответственно первой и второй регрессий;

$$Z_1 = \begin{pmatrix} X^T \cdot X & X^T \cdot \hat{Y}_1 \\ \hat{Y}_1^T \cdot X & \hat{Y}_1^T \cdot \hat{Y}_1 \end{pmatrix}, \; Z_2 = \begin{pmatrix} X^T \cdot X & X^T \cdot \hat{Y}_2 \\ \hat{Y}_2^T \cdot X & \hat{Y}_2^T \cdot \hat{Y}_2 \end{pmatrix} \text{ – блочные матрицы;} \tag{8.4}$$

$$\hat{Y}_1 = \begin{pmatrix} \hat{y}_{21} \\ \hat{y}_{22} \\ \vdots \\ \hat{y}_{2n} \end{pmatrix}, \quad \hat{Y}_2 = \begin{pmatrix} \hat{y}_{11} \\ \hat{y}_{12} \\ \vdots \\ \hat{y}_{1n} \end{pmatrix}$$ – векторы-столбцы, полученные из матрицы $\hat{Y}$ путем вычеркивания соответственно первого столбца и второго столбца, если определитель $\det[Z] \neq 0$.

8.2 Алгоритм построения и анализа эконометрической модели методом 2МНК в оболочке электронных таблиц MS Excel

1 Идентификация переменных и спецификация модели.

2 Проверка условий идентифицированности для каждого уравнения структурной формы модели.

3 Составление прогнозной формы модели и оценка ее параметров 1МНК (первый шаг 2МНК).

4 Определение оценок параметров структурной формы модели (второй шаг 2МНК).

5 Оценка адекватности структурной формы модели.

6 Прогнозирование по оцененной прогнозной модели.

7 Экономический анализ по оцененной структурной модели.

8.3 Реализация двухшагового метода наименьших квадратов для решения СОУ в оболочке электронных таблиц MS Excel

На основе статистических данных за 10 лет об экспорте, импорте, национальном доходе и среднем обороте внешней торговли стран, с которыми поддерживаются внешнеэкономические отношения, оценить параметры прогнозной и структурной форм эконометрической модели «экспорт – импорт» блока внешней торговли *N*-й страны (таблица 7.1, в усл. ден. ед.), допуская, что структурная форма модели имеет вид (8.1), т. е.

$$\begin{cases} Y_1 = b_{12}Y_2 + a_{10} + a_{11}X_1 + a_{12}X_2 + e_1, \\ Y_2 = b_{21}Y_1 + a_{20} + a_{12}X_1 + a_{22}X_2 + e_2. \end{cases}$$

Поставленную задачу решим в среде Excel по следующему алгоритму.

8.3.1 Идентификация переменных и спецификация модели. Блок исходных данных формируется в массиве C3:F12, в ячейки B3:B12 вводятся номера наблюдений (рисунок 8.1).

Пусть векторы эндогенных переменных:

Y_1 – экспорт;

Y_2 – импорт

и экзогенных переменных:

X_1 – национальный доход страны;

X_2 – внешний товарооборот страны, с которыми поддерживаются внешнеэкономические отношения.

	A	B	C	D	E	F	G	H
1								
2		№ п/п	X1	X2	Y1	Y2	Матрица оценок	
3		1	48	4,5	50	65	48,3648	64,2602
4		2	55	5,1	63	79	61,4136	76,0805
5		3	61	5,6	72	85	72,3698	86,0124
6		4	65	6,3	81	97	85,5383	97,7612
7		5	69	6,9	96	106	97,1074	108,1116
8		6	72	7,2	105	111	103,3852	113,7768
9		7	76	7,8	116	125	114,9543	124,1272
10		8	80	8,6	129	137	129,7221	137,2744
11		9	86	8,9	130	142	137,4796	144,4094
12		10	88	9,1	150	153	141,6648	148,1862
13		Сумм	700	70	992	1100	992	1100

Рисунок 8.1 – Исходные данные. Матрица оценок эндогенных переменных по прогнозной модели

8.3.2 Проверка условия на идентифицированность для каждого уравнения структурной формы модели. Модель «экспорт – импорт» имеет вид (8.1), т. е. является сверхидентифицированной, поэтому для оценки ее параметров используем 2МНК.

8.3.3 Составление прогнозной формы модели и оценка ее параметров 1МНК (первый шаг 2МНК). Прогнозная форма модели имеет вид (7.5), т. е. $Y = X \cdot C + U$, оценку параметров этой модели получим по методу 1МНК, т. е. по формуле (7.8): $\hat{C} = \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y$.

Составим в диапазоне **C17:E26** матрицу *X*, в диапазоне G17:H26 матрицу *Y*, копируем исходные данные (рисунок 8.2). Используя встроенную функцию **ТРАНСП**, в диапазоне **C28:L30** вычислим транспонированную матрицу *X* и далее по алгоритму вычисления параметров модели, который приведен в практической работе №3, в диапазоне **C32:D34** получим матрицу *C* оценок параметров прогнозной формы модели. Для проверки правильности определения оценок параметров модели и регрессионных статистик в оболочке электронных таблиц Excel используем дважды (для каждой регрессии) пакет **Анализ дан-**

ных. Откроем новые листы для **Вывода итогов**, введем имя новых листов «Прогноз регрессия 1» и «Прогноз регрессия 2».

Для сравнения полученных значений параметров модели по формулам и при помощи **Пакета анализа** скопируем с листов «Прогноз регрессия 1» и «Прогноз регрессия 2» значения ячеек столбца ***Коэффициенты*** строк ***Y*-пересечение**, X_1, X_2 в ячейки **I33:L35**. Полученные значения совпали.

Составим оцененную прогнозную модель в диапазоне C37:H38, исходя из формулы (7.6).

Таким образом, оцененная прогнозная модель имеет вид

$$\begin{cases} \hat{Y}_1 = -47{,}282 + 0{,}493X_1 + 15{,}994X_2, \\ \hat{Y}_2 = 22{,}187 + 0{,}490X_1 + 15{,}984X_2. \end{cases}$$

	B	C	D	E	F	G	H	I	J	K	L
16											
17		1	48	4,5		50	65				
18		1	55	5,1		63	79				
19		1	61	5,6		72	85				
20		1	65	6,3		81	97				
21	X=	1	69	6,9	У=	96	106				
22		1	72	7,2		105	111				
23		1	76	7,8		116	125				
24		1	80	8,6		129	137				
25		1	86	8,9		130	142				
26		1	88	9,1		150	153				
27											
28		1	1	1	1	1	1	1	1	1	1
29	XT=	48	55	61	65	69	72	76	80	86	88
30		4,5	5,1	5,6	6,3	6,9	7,2	7,8	8,6	8,9	9,1
31									Регрессия		
32					-47,282	-22,187		1-я регрессия		2-я регрессия	
33				C=	0,493	0,490		Y-пересечение	-47,282	Y-пересечение	-22,187
34					15,994	13,984		X 1	0,493	X 1	0,490
35	Прогнозная форма модели							X 2	15,994	X 2	13,984
37	Y1=	-47,282	+	0,493	*X1+	15,994	*X2				
38	Y2=	-22,187	+	0,490	*X1+	13,984	*X2				

Рисунок 8.2 – Определение оценок параметров прогнозной формы модели

8.3.4 Определение оценок параметров структурной формы модели (второй шаг 2МНК). Оценки параметров структурной формы модели получим на основе формулы (8.3), выполнив следующие действия.

1 Найдем матрицу расчетных значений эндогенных переменных по формуле (7.7) $\hat{Y} = X \cdot \hat{C}$ в диапазоне G3:H12, используя встроенную математическую функцию МУМНОЖ.

2 Составим в диапазоне C40:C49 (рисунок 8.3) вектор $\hat{Y}_1$, который получается с матрицы $\hat{Y}$ вычеркиванием соответствующего i-го столбца. Для этого используем ссылки на диапазон H3:H12 (см. рисунок 8.2).

3 Используя встроенные математические функции ТРАНСП и МУМНОЖ, определим матрицы $\hat{Y}^T$, $X \cdot \hat{Y}_1$, $\hat{Y}_1^T \cdot X$, $\hat{Y}_1^T \cdot \hat{Y}_1$, $X^T \cdot Y_1$, $\hat{Y}_1^T \cdot Y_1$ (см. рисунок 8.3),

которые являются составляющими блочных матриц $Z_1 = \begin{pmatrix} X^T \cdot X & X^T \cdot \hat{Y}_1 \\ \hat{Y}_1^T \cdot X & \hat{Y}_1^T \cdot \hat{Y}_2 \end{pmatrix}$ и $\begin{pmatrix} X^T \cdot Y_1 \\ \hat{Y}_1^T \cdot Y_1 \end{pmatrix}$.

4 Составим блочную матрицу Z_1 в диапазоне C53:F56, используя ссылки на соответствующие диапазоны. Аналогичным образом составим блочную матрицу $\begin{pmatrix} X^T \cdot Y_1 \\ \hat{Y}_1^T \cdot Y_1 \end{pmatrix}$ в диапазоне H46:H49.

	B	C	D	E	F	G	H	I	J	K
39										
40		64,260		1100						
41		76,081	XTY1=	80397						
42		86,012		8122,4						
43		97,761								
44	Y1=	108,112	У1TX=	1100	80397	8122,4				
45		113,777								
46		124,127					992			
47		137,274	У1TУ1=	128571,34		XTУ1=	73222			
48		144,409					7414,4			
49		148,186					117551,22			
50	Y1=									
51	64,260	76,081	86,012	97,761	108,112	113,777	124,127	137,274	144,409	148,186
52										
53		10	700	70	1100		-4,567E+13	1,00851E+12	2,878E+13	-2,05832E+12
54	Z1=	700	50536	5089,1	80397	Z-11=	1,009E+12	-22271516489	-6,357E+11	45454957818
55		70	5089,1	513,58	8122,4		2,878E+13	-6,3565E+11	-1,814E+13	1,29733E+12
56		1100	80397	8122,4	128571,34		-2,058E+12	45454957818	1,297E+12	-92771104797
57	detZ1=	-4,959E-08								
58		1	0,015625	-0,25	-0,015625		-2	-224	-22	-320
59	AA-1=	32	1,5	0	-1	A-1A=	0,03125	3,5	0,3125	4
60		0	0,125	0	0		0	0	0	0
61		32	1	0	0		0,015625	0	0,125	2

Рисунок 8.3 – Определение оценок параметров структурной формы первой регрессии

5 Вычислим определитель матрицы Z_1 в ячейке C57, используя встроенную функцию МОПРЕД. Получили $\det Z_1 = -4{,}959E-08$, т. е. определитель матрицы практически равен нулю, это означает, что обратной матрицы не существует.

6 Вычислим обратную матрицу Z_1^{-1} в диапазоне H53:K56, используя функцию МОБР.

7 Для проверки вычислений воспользуемся определением обратной матрицы, т. е. равенством $A^{-1}A = AA^{-1} = 1$. Оба произведения не являются единичными матрицами, поэтому приходим к выводу, что обратной матрицы не существует и применить матричный метод решения (8.3) невозможно.

8 Составим систему уравнений, соответствующую матричному уравнению $Z_1\hat{d}_1 = \begin{pmatrix} X^T \cdot Y_1 \\ \hat{Y}_1^T \cdot Y_1 \end{pmatrix}$ (согласно формуле (8.3)), и решим ее методом Жордана – Гаусса (рисунок 8.4). В третьей таблице Гаусса получили нулевой порядок, а это свидетельствует о том, что система уравнений имеет множество решений.

Таким образом, можно составить множество эконометрических моделей по приведенным статистическим данным, но их экономическая интерпретация не имеет значения, поэтому для оценки параметров структурной формы эконометрической модели «экспорт – импорт» блока внешней торговли N-й страны по приведенным данным необходимо использовать другой метод.

	B	C	D	E	F	G
62	№ итерации	z11	z12	z13	z14	Свободные члены
63	исходная	10	700	70	1100	992
64	матрица	700	50536	5089,1	80397	73222
65		70	5089,1	513,58	8122,4	7414,4
66		1100	80397	8122,4	128571,34	117551,22
67	I	1	70	7	110	99,2
68		0	**1536**	189,1	3397	3782
69		0	189,1	23,58	422,4	470,4
70		0	3397	422,4	7571,3405	8431,219
71	II	1	0	-1,6178	-44,8112	-73,1568
72		0	1	0,1231	2,2116	2,4622
73		0	0	**0,2995**	4,1886	4,7905
74		0	0	4,1886	58,5742	66,9911
75	III	1	0	0	-22,2	-47,3
76		0	1	0	0,5	0,5
77		0	0	1	13,9842	15,9937
78		0	0	0	0	0

Рисунок 8.4 – Решение системы уравнений методом Жордана – Гаусса

8.4 Модифицированный двухшаговый метод наименьших квадратов (М2МНК)

Рассмотренный 2МНК имеет недостаток: предопределенные значения $\hat{Y}_i$ получаются с прогнозной формы, отклонения которых являются преобразованными по отношению к отношениям структурной формы модели.

Рассмотрим сверхидентифицированную систему двух взаимозависимых линейных регрессий (8.1), т. е.

$$\begin{cases} Y_1 = b_{12}Y_2 + a_{10} + a_{11}X_1 + a_{12}X_2 + e_1, \\ Y_2 = b_{21}Y_1 + a_{20} + a_{12}X_1 + a_{22}X_2 + e_2. \end{cases}$$

Алгоритм модифицированного 2МНК состоит из двух шагов:

- первый шаг состоит в оценке 1МНК параметров прогнозной формы модели (7.6);

- второй шаг заключается в оценке 1МНК параметров при эндогенных величинах модели в структурной форме со значениями, которые представляют собой отклонения эндогенных значений, полученных на первом шаге. На основе оцененных параметров двух матриц *C* и *B* определяются элементы третьей матрицы *A*.

Алгоритм модифицированного 2МНК

Первый шаг М2МНК.

1 Проверить каждое уравнение структурной формы модели на идентифицированность. Если система сверхидентифицирована, то переходим к следующему пункту.

2 Записать прогнозную форму модели в матричном виде, т. е. модель (7.5)

$$Y = X \cdot C + U,$$

где $C = \begin{pmatrix} c_{10} & c_{20} \\ c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix}$ – матрица коэффициентов прогнозной формы модели, допуская, что определитель матрицы $B = \begin{pmatrix} -1 & b_{21} \\ b_{12} & 1 \end{pmatrix}$ не равен нулю.

3 Оценить параметры прогнозной формы модели методом 1МНК, т. е. по формуле (7.8):

$$\hat{C} = \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y,$$

если определитель $\det\left[X^T X\right] \neq 0$.

Второй шаг М2МНК.

1 Определить матрицу оценок эндогенных переменных, используя оцененную прогнозную модель:

$$\hat{Y} = X \cdot \hat{C} = \begin{pmatrix} \hat{y}_{11} & \hat{y}_{21} \\ \hat{y}_{12} & \hat{y}_{22} \\ \vdots & \vdots \\ \hat{y}_{1n} & \hat{y}_{12} \end{pmatrix}.$$

2 Составить матрицу отклонений прогнозной формы модели:

$$U = Y \cdot \hat{Y} = \begin{pmatrix} y_{11} - \hat{y}_{11} & y_{21} - \hat{y}_{21} \\ y_{12} - \hat{y}_{12} & y_{22} - \hat{y}_{22} \\ \vdots & \vdots \\ y_{1n} - \hat{y}_{1n} & y_{2n} - \hat{y}_{2n} \end{pmatrix}. \tag{8.5}$$

3 Оценить параметры системы регрессий между отклонениями эндогенных значений

$$\begin{cases} U_1 = b_{12} U_2 + e_1, \\ U_2 = b_{22} U_1 + e_2. \end{cases} \tag{8.6}$$

методом 1МНК, т. е. по формулам

$$\hat{b}_{12} = \frac{\sum_{i=1}^{n} u_{1i} u_{2i}}{\sum_{i=1}^{n} u_{2i}^2}, \qquad \hat{b}_{21} = \frac{\sum_{i=1}^{n} u_{1i} u_{2i}}{\sum_{i=1}^{1n} u_{1i}^2}. \tag{8.7}$$

4 Составить матрицу *B* и определить оценки матрицы *A*, используя матричное уравнение (7.9), т. е.

$$\hat{A} = \hat{C} \cdot \hat{B}.$$

8.5 Реализация модифицированного двухшагового метода наименьших квадратов для решения СОУ в оболочке электронных таблиц MS Excel

Оценить параметры прогнозной и структурной форм эконометрической модели «экспорт – импорт» (задача подраздела 8.3), используя модифицированный 2МНК. Найти значения прогноза экспорта и импорта, если прогнозные значения национального дохода и среднего оборота внешней торговли страны равны соответственно 90 и 10 усл. ден. ед.

Поставленную задачу решим в MS Excel по следующему алгоритму.

Алгоритм построения и анализа эконометрической модели модифицированным методом 2МНК в оболочке электронных таблиц MS Excel:

1 Идентификация переменных и спецификация модели.

2 Проверка условий идентифицированности для каждого уравнения структурной формы модели.

3 Составление прогнозной формы модели и оценка ее параметров 1МНК (первый шаг М2МНК).

4 Определение оценок параметров структурной формы модели (второй шаг М2МНК).

5 Оценка адекватности структурной формы.

6 Прогнозирование по оцененной прогнозной модели.

7 Экономический анализ по оцененной структурной модели.

Решение

8.5.1 Идентификация переменных и спецификация модели рассмотрены в подразделе 8.3.

8.5.2 Проверка условий идентифицированности для каждого уравнения структурной формы модели рассмотрена в подразделе 8.3.

8.5.3 Составление прогнозной формы модели и оценка ее параметров 1МНК (первый шаг М2МНК). Первый шаг модифицированного 2МНК совпадает с первым шагом 2МНК, поэтому решение примера начнем со второго шага, используя прогнозную модель

$$\begin{cases} \hat{Y}_1 = -47{,}282 + 0{,}493X_1 + 15{,}994X_2, \\ \hat{Y}_2 = 22{,}187 + 0{,}490X_1 + 15{,}984X_2. \end{cases}$$

8.5.4 Определение оценок параметров структурной формы модели (второй шаг М2МНК). Оценку параметров структурной формы модели получим, используя матрицу оценок эндогенных переменных, которая расположена в диапазоне G3:H12 (см. рисунок 8.1), выполнив следующие действия:

1 Вычислим элементы матрицы U отклонений прогнозной формы модели по формуле (8.5) следующим образом: в ячейку K3 введем формулу =E3-G3, скопируем ее в ячейку L3 и далее – в остальные ячейки диапазона K3:L12 (рисунок 8.5).

2 Вычислим суммы квадратов отклонений первой и второй эндогенных переменных соответственно в ячейках K14 и L14, используя встроенную функцию СУММКВ.

3 Вычислим сумму произведений соответствующих отклонений первой и второй эндогенных переменных в ячейке L15, используя встроенную функцию =СУММПРОИЗВ(K3:K12;L3:L12).

4 Вычислим оценки параметров структурной формы модели $\hat{b}_{12}$ и $\hat{b}_{21}$ по формулам (8.7) соответственно в ячейках U2 и W2 (рисунок 8.6), в ячейку U2 вводим формулу =L15/L14, а в ячейку W2 – формулу =L15/K14.

5 Составим матрицу в диапазоне U4:V5, используя ссылки на ячейки U2 и W2.

6 Вычислим произведение матриц CD в диапазоне X3:Y5, используя встроенную функцию =МУМНОЖ(F32:G34;U4:V5).

7 Составим матрицу $A = -CD$ в диапазоне AA3:AB5.

8 Составим оцененную структурную форму модели в диапазоне T7:AC8 по формуле (8.1), используя ссылки на соответствующие элементы матрицы (см. рисунок 8.6).

	J	K	L	M	N	O	P	Q	R
1									
2	№ п/п	Отклонения u1	Отклонения u2	Структурная y2	Структурная y2	Отклонения e1	Отклонения e2	y1-yср1	y1-yср2
3	1	1,6352	0,7398	49,3030	64,9567	0,6970	0,0433	-49,8970	-45,0433
4	2	1,5864	2,9195	65,1160	76,7562	-2,1160	2,2438	-34,0840	-33,2438
5	3	-0,3698	-1,0124	71,0859	85,8549	0,9141	-0,8549	-28,1141	-24,1451
6	4	-4,5383	-0,7612	84,5729	95,8283	-3,5729	1,1717	-14,6271	-14,1717
7	5	-1,1074	-2,1116	94,4295	107,6399	1,5705	-1,6399	-4,7705	-2,3601
8	6	1,6148	-2,7768	99,8638	114,4646	5,1362	-3,4646	0,6638	4,4646
9	7	1,0457	0,8728	116,0612	124,5725	-0,0612	0,4275	16,8612	14,5725
10	8	-0,7221	-0,2744	129,3742	136,9668	-0,3742	0,0332	30,1742	26,9668
11	9	-7,4796	-2,4094	134,4240	141,2237	-4,4240	0,7763	35,2240	31,2237
12	10	8,3352	4,8138	147,7695	151,7363	2,2305	1,2637	48,5695	41,7363
13	Сумм			992,0000	1100,0000				
14	Суммкв	156,7924	52,6595			72,1030	24,2161	9473,4970	7599,7839
15	Суммпроизв		66,7810						

Рисунок 8.5 – Расчет отклонений по прогнозной и структурной формам моделей

	T	U	V	W	X	Y	Z	AA	AB	AC
1										
2	b12=	1,268	b21=	0,426						
3					19,145	2,049		-19,145	-2,049	
4	B=	-1	0,426	CB=	0,128	-0,280	A=	-0,128	0,280	
5		1,268	-1		1,741	-7,172		-1,741	7,172	
6			Структурная форма модели имеет вид							
7	Y1=	1,268	*Y2+	-19,145	+	-0,128	*X1	+	-1,741	*X2
8	Y2=	0,426	*Y1+	-2,049	+	0,280	*X1	+	7,172	*X2
9				Оценка адекватности модели						
10	1 регрессия	R-квадрат	0,9924	Fрасч	260,77	Fтабл(3,6)	4,76	Модель адекватна		
11	2 регрессия	R-квадрат	0,9968		625,663			Модель адекватна		

Рисунок 8.6 – Определение оценок параметров структурной формы модели по М2МНК. Оценка адекватности модели

8.5.5 Оценку адекватности структурной формы эконометрической модели проведем стандартным образом по коэффициенту детерминации и критерию Фишера.

1 В диапазоне M3:M12 (см. рисунок 8.5) вычислим оценки первой эндогенной переменной по структурному уравнению, а в диапазоне L3:L12 вычислим оценки второй эндогенной переменной.

2 В ячейках M13 и L13, используя кнопку автосуммы, найдем соответствующие суммы оценок эндогенных переменных. Полученные суммы совпали с соответствующими исходными значениями и теми, которые получили по прогнозной модели.

3 В диапазоне O3:P12 вычислим матрицу *e* отклонений структурной модели аналогично определению матрицы отклонений *U* прогнозной модели. В ячейках O13 и P13 найдем суммы квадратов отклонений.

4 В диапазоне Q3:R12 рассчитаем отклонения от среднего для первой и второй переменных: в ячейке Q3 вводим формулу =M3-СРЗНАЧ($E$3:$E$12) и копируем ее для остальных ячеек диапазона.

5 В диапазоне O14:R14 вычислим суммы квадратов соответствующих отклонений, используя встроенную функцию СУММКВ.

6 В ячейках V10 и V11 вычислим коэффициенты детерминации для каждой модели.

7 В ячейках X10 и X11 вычислим расчетные значения критерия Фишера для каждой модели.

8 Критическое значение *F*-критерия находим по статистическим таблицам *F*-распределения Фишера для соответствующих степеней свободы $k_1 = 3$ и $k_2 = 6$ и доверительного интервала $P = 0{,}95$.

Расчетные значения *F*-распределения значительно больше критических, поэтому с надежностью $P = 0{,}95$ можно считать, что принятая модель «экспорт-импорт» адекватна статистическим данным и на основе этой модели можно проводить экономический анализ. Следовательно, модифицированный 2VYR дает возможность получить адекватную эконометрическую модель. Для эконометрического анализа и прогнозирования развития экономических отношений следует оценить параметры структурной системы регрессий модифицированным 2МНК.

8.5.6 Прогнозирование по оцененной прогнозной модели. Для нахождения точечных оценок прогноза используется система регрессий в прогнозной форме. По условию примера прогнозные значения экзогенных переменных равны $x_{1p} = 90$, $x_{2p} = 10$. Получили прогнозные значения эндогенных переменных, т. е. прогнозное значение экспорта – 157 усл. ден. ед., а импорта – 161,8 усл. ден. ед.

8.5.7 Экономический анализ по оцененной структурной модели. Анализ структурной формы модели приводит к следующим результатам:

- при увеличении Y_2 (импорта) на единицу, при неизменных прочих значениях Y_1 (экспорт) увеличится на 1,3 усл. ден. ед.;

- при увеличении X_1 (национального дохода) на единицу при неизменных других переменных Y_1 (экспорт) уменьшится на 0,1 усл. ден. ед.;

- при увеличении X_2 (среднего внешнего товарооборота стран блока) на единицу, при неизменных прочих значениях Y_1 (экспорт) уменьшится на 1,7 усл. ден. ед.;

- при увеличении Y_1 (экспорта) на единицу при неизменных прочих значениях Y_2 (импорт) увеличится на 0,4 усл. ден. ед.;

- при увеличении X_1 (национального дохода) на единицу при неизменных прочих значениях Y_2 (импорт) увеличится на 0,3 усл. ден. ед.;

- при увеличении X_2 (среднего внешнего товарооборота стран блока) на единицу при неизменных прочих значениях Y_2 (импорт) увеличится на 7,2 усл. ден. ед.

8.6 Варианты заданий

Итоговыми показателями эффективности работы строительных организаций является общая реализация строительной продукции (объем сданных строительных работ) и общая рентабельность. На основе статистических данных по 12 строительно-монтажным организациям об общей реализации строительной продукции (Y_1), общей рентабельности (Y_2), доле сборных конструкций, деталей и изделий в себестоимости строительно-монтажных работ (X_1), выработке на одного рабочего (X_2) оценить параметры прогнозной и структурной форм эконометрической модели, допуская, что структурная форма модели имеет вид

$$\begin{cases} Y_1 = b_{12}Y_2 + a_{10} + a_{11}X_1 + a_{12}X_2 + e_1, \\ Y_2 = b_{21}Y_1 + a_{20} + a_{12}X_1 + a_{22}X_2 + e_2. \end{cases}$$

Руководство объединения хотело бы знать, насколько изменятся общая реализация строительной продукции и общая рентабельность, если часть сборных конструкций, деталей и изделий в себестоимости строительно-монтажных работ и выработка на одного рабочего составят 95 % от соответствующих значений 12-го СМУ.

Для выполнения этого задания используются исходные данные для задания по КМНК.

Контрольные вопросы

1 Назовите признаки нормального распределения.
2 Что такое медиана?
3 Что такое регрессия?
4 Что показывает коэффициент парной корреляции?
5 Что показывает коэффициент множественной детерминации?
6 Для чего необходим *t*-критерий Стьюдента?
7 Для чего необходим *F*-критерий Фишера?
8 Что такое гетероскедастичность?
9 Какие тесты для обнаружения гетероскедастичности вы знаете?
10 Назовите последствия гетероскедастичности.
11 Что такое мультиколлинеарность?
12 Какие причины возникновения мультиколлинеарности вы знаете?
13 Как выявить наличие мультиколлинеарности?
14 Что такое автокорреляция?
15 Какие вы знаете причины автокорреляции?
16 Какие последствия автокорреляции вы знаете?
17 Для решения каких систем одновременных уравнений предназначен КМНК?
18 Каков алгоритм КМНК?
19 Для решения каких систем одновременных уравнений предназначен 2МНК?
20 Каков алгоритм 2МНК?

Литература

1 Якимова, Л. П. Эконометрия: практикум с использованием компьютера : учеб. пособие для студентов экономических специальностей высших учебных заведений / Л. П. Якимова. – Алчевск : ДГМИ, 2007. – 189 с.

2 Копернак, Ю. Н. Лабораторный практикум по математической статистике : учеб. пособие для студентов экономических специальностей ДГМИ / Ю. Н. Копернак. – Алчевск : ДГМИ, 2005. – 90 с.

3 Алехина, А. Э. Эконометрика : учеб.-метод. пособие / А. Э. Алехина, С. А. Поттосина. – Минск : БГУИР, 2011. – 90 с.

4 Носко, В. П. Эконометрика для начинающих (Дополнительные главы) / В. П. Носко. – М. : ИЭПП, 2005. – С. 379.

5 Эконометрика / сост. Л. Н. Леванова. – Саратов, 2003.

6 Эконометрия / В. И. Суслов [и др.] – Новосибирск : Издательство СО РАН, 2005. – 744 с.

7 Бородич, С. А. Вводный курс эконометрики : учеб. пособие / С. А. Бородич. – Минск : БГУ, 2000. – 354 с.

ПРИЛОЖЕНИЕ А
(справочное)

Справочные таблицы

Таблица А.1 – Плотность вероятности нормального распределения $f(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$

t	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3856	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3696
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3148	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3189	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2372	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1624	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0718	0734	0721	0707	0694	0681	0669
1,9	0656	0614	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0289	0277	0270	0261	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0031.
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018

t	0	1	2	3	4	5	6	7	8	9
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Таблица А.2 – *F*-распределение для уровня значимости $\alpha = 0{,}05$

k_1	k_1										
	1	2	3	4	5	6	7	8	9	10	11
1	161	200	216	225	230	234	237	239	241	242	243
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,40
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,76
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45
20	4,35	3,49	3,10	2,87	2,71	2,60	2,,52	2,45	2,40	2,35	2,31
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,22
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	2,12
40	4,08	3,23	2,84	2,61	2,45	2,34	,2,25	2,18	2,12	2,07	2,04
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,98
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,01	1,97	1,93
100	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,88
200	3,89	3,04	2,65	2,41	2,26	2,14	2,05	1,98	1,92	1,87	1,83
400	3,86	3,02	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,81
∞	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88	1,83	1,79

Продолжение таблицы А.2

k_1	k_1											
	12	13	16	20	30	40	60	75	100	200	500	∞
1	244	245	246	248	250	251	252	253	253	254	254	254
2	19,41	19,42	19,43	19,44	19,46	19,47	19,47	19,48	19,49	19,49	19,50	19,50
3	8,74	8,71	8,69	8,66	8,62	8,60	8,58	8,57	8,56	8,54	8,54	8,53
4	5,91	5,87	5,84	5,80	5,74	5,71	5,70	5,68	5,66	5,65	5,64	5,63
5	4,68	4,64	4,60	4,56	4,50	4,46	4,44	4,42	4,40	4,38	4,38	4,36
6	4,00	3,96	3,92	3,87	3,81	3,77	3,75	3,72	3,71	3,69	3,68	3,67
7	3,57	3,52	3,49	3,44	3,38	3,34	3,32	3,29	3,28	3,25	3,24	3,23
8	3,28	3,23	3,20	3,15	3,08	3,05	3,03	3,00	2,98	2,96	2,94	2,93
9	3,07	3,02	2,98	2,93	2,86	2,82	2,80	2,77	2,76	2,73	2,72	2,71
10	2,91	2,86	2,82	2,77	2,70	2,67	2,64	2,61	2,59	2,56	2,55	2,54
11	2,79	2,74	2,70	2,65	2,57	2,53	2,50	2,47	2,45	2,42	2,41	2,40
12	2,69	2,64	2,60	2,54	2,46	2,42	2,40	2,36	2,35	2,32	2,31	2,30
14	2,53	2,48	2,44	2,39	2,31	2,27	2,24	2,21	2,19	2,16	2,14	2,13
16	2,42	2,37	2,33	2,28	2,20	2,16	2,13	2,09	2,07	2,04	2,02	2,01
20	2,28	2,23	2,18	2,12	2,04	1,99	1,96	1,92	1,90	1,87	1,85	1,84
24	2,18	2,13	2,09	2,02	1,94	1,89	1,86	1,82	1,80	1,76	1,74	1,73
30	2,09	2,04	1,99	1,93	1,84	1,79	1,76	1,72	1,69	1,66	1,64	1,62
40	2,00	1,95	1,90	1,84	1,74	1,69	1,66	1,61	1,59	1,55	1,53	1,51
50	1,95	1,90	1,85	1,78	1,69	1,63	1,60	1,55	1,52	1,48	1,46	1,44
70	1,89	1,84	1,79	1,72	1,62	1,56	1,53	1,47	1,45	1,40	1,37	1,35
100	1,85	1,79	1,75	1,68	1,57	1,51	1,48	1,42	1,39	1,34	1,30	1,28
200	1,80	1,74	1,69	1,62	1,52	1,45	1,42	1,35	1,32	1,26	1,22	1,19
400	1,78	1,72	1,67	1,60	1,49	1,42	1,38	1,32	1,28	1,22	1,16	1,13
∞	1,75	1,69	1,64	1,57	1,46	1,40	1,35	1,28	1,24	1,17	1,11	1,00

Таблица А.3 – Значения t-распределения Стьюдента

k	P			k	P		
	0,95	0,99	0,999		0,95	0,99	0,999
4	2,78	4,60	8,61	19	2,093	2,861	3,883
5	2,57	4,03	6,86	20	2,088	2,845	3,849
6	2,45	3,71	5,96	25	2,064	2,797	3,745
7	2,37	3,50	5,41	30	2,045	2,756	3,659
8	2,31	3,36	5,04	35	2,032	2,729	3,600
9	2,26	3,25	4,78	40	2,023	2,708	3,558
10	2,23	3,17	4,59	45	2,016	2,692	3,527
11	2,20	3,11	4,44	50	2,009	2,679	3,502
12	2,18	3,06	4,32	60	2,001	2,662	3,464
13	2,16	3,01	4,22	70	1,996	2,649	3,439
14	2,15	2,98	4,14	80	1,991	2,640	3,418
15	2,13	2,95	4,07	90	1,987	2,633	3,403
16	2,12	2,92	4,02	100	1,984	2,627	3,392
17	2,11	2,90	3,97	120	1,980	2,617	3,374
18	2,10	2,88	3,92	∞	1,960	2,576	3,291

Таблица А.4 – Значения χ^2 в зависимости от вероятности $P\,(\chi^2 > \chi^2_\alpha)$ и числа степени свободы k

Число степеней свободы k	Вероятность															
	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,005	0,002	0,001
1	0,00016	0,0006	0,0039	0,016	0,064	0,148	0,455	1,07	1,64	2,7	3,8	5,4	6,6	7,9	9,5	10,83
2	0,020	0,040	0,103	0,211	0,446	0,713	1,386	2,41	3,22	4,6,	6,0	7,8	9,2	11,6	12,4	13,8
3	0,115	0,185	0,352	0,584	1,005	1,424	2,366	3,66	4,64	6,3	7,8	9,8	!1,3	12,8	14,8	16,3
4	0,30	0,43	0,71	1,06	1,65	2,19	3,36	4,9	6,0	7,8	9,5	11,7	13,3	14,9	16,9	18,5
5	0,55	0,75	1,14	1,61	2,34	3,00	4,35	6,1	7,3	9,2	11,1	13,4,	15,1	16,3	18,9	20,5
6	0,187	1,13	1,63	2,20	3,07	3,83	5,35	7,2	8,6	10,6	12,6	15,0	16,8	18,6	20,7	22,5
7	1,24	1,56	2,17	2,83	3,82	4,67	6,34	8,4	9,8	12,0	14,1	16,6	18,5	20,3	22,6	24,3
8	1,65	2,03	2,73	3,49	4,59	5,53	7,34	9,5	11,0	13,4	15,5	18,2	20,1	21,9	24,3	21,6
9	2,09	2,53	3,32	4,17	5,38	6,39	8,35	10,7	12,2	14,7	16,9	19,7	21,7	23,6	26,1	27,9
10	2,56	3,06	3,94	4,86	6,18	7,27	9,34	11,8	13,4,	16,0	18,3	21,2	23,2	25,2	27,7	29,6
11	3,1	3,6	4,6	5,6	7,0	8,1	10,3	12,9	14,6	17,3	19,7	22,6	24,7	26,8	29,4	31,3
12	3,6	4,2	5,2	6,3	7,8	9,0	11,3	14,0	15,8	18,5	21,0	24,1	26,2	28,3	31,0	32,9
13	4,1	4,8	5,9	7,0	8,6	9,9	12,3	15,1	17,0	,19,8	22,4	25,5	27,7	29,8	32,5	34,5
14	4,7	5,4	6,6	7,8	9,5	10,8	13,3	16,2	18,2	21,1	23,7	26,9	29,1	31,0	34,0	36,1
15	5,2	6,0	7,3	8,5	10,3	11,7	14,3	17,3	19,3	22,3	25,0	28,3	30,6	32,5	35,5	37,7
16	5,8	6,6	8,0	9,3	11,2	12,6	15,3	18,4	20,5	23,5	26,3	29,6	32,0	34,0	37,0	39,2

Число степеней свободы k	Вероятность															
	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,05	0,002	0,001
17	6,4	7,3	8,7	10,1	12,0	13,5	16,3	19,5	21,6	24,8	27,6	31,0	33,4	35,5	38,5	40,8
18	7,0	7,9	9,4	10,9	12,9	14,4	17,3	20,6	22,8	26,0	28,9	32,3	34,8	37,0	40,0	42,3
19	7,6	8,6	10,!	11,7	13,7	15,4	18,3	21,7	23,9	27,2	30,1	33,7	36,2	38,5	41,5	43,8
20	8,3	9,2	10,9	12,4	14,6	16,3	19,3	22,8	25,0	28,4	31,4	35,0	37,6	40,0	43,0	45,3
21	8,9	9,9	11,6	13,2	15,4	17,2	20,3	23,9	26,2	29,6	32,7	36,3	38,9	41,5	44,5	46,8
22	9,5	10,6	12,3	14,0	16,3	18,1	21,3	24,9	27,3	30,8	33,9	37,7	40,3	42,5	46,0	48,3
23	10,2	11,3	13,1	14,8	17,2	19,0	22,3	26,0	28,4	32,0	35,2	39,0	41,6	44,0	47,5	49,7
24	10,9	12,0	13,8	15,7	18,1	19,9	23,3	27,1	29,6	33,2	36,4,	40,3	43,0	45,5	48,5	51,2
25	11,5	12,7	14,6	16,5	18,9	20,9	24,3	28,1	30,7	34,4	37,7	41,6	44,3	47,0	50,0	52,6
26	12,2	13,4	15,4	17,3	19,8	21,8	25,3	29,3	31,8	35,6	38,9	42,9	45,6	48,0	51,5	54,1
27	12,9	14,1	16,2	18,1	20,7	22,7	26,3	30,3	32,9	36,7	40,1	44,1	47,0	49,5	53,0	55,5
28	13,6	14,8	16,9	18,9	21,6	23,6	27,3	31,4	34,0	37,9	41,3	45,4	48,3	51,0	54,5	56,9
29	14,3	15,6	17,7	19,8	22,5	24,6	28,3	32,5	35,1	39,1	42,6	46,7	49,6	52,5	56,0	58,3
30	15,0	16,3	18,5	20,6	23,4	25,5	29,3	33,5	36,3	40,3	43,8	48,0	50,9	54,0	57,5	59,7

Св. план 2018, поз. 62

Учебное издание

ЭКОНОМЕТРИКА И ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ И МОДЕЛИ

В двух частях
Часть 1

Голда Ольга Алексеевна
Матвейчук Наталья Михайловна

ЭКОНОМЕТРИКА

УЧЕБНО-МЕТОДИЧЕСКОЕ ПОСОБИЕ

Редактор *Е. И. Костина*
Корректор *Е. Н. Батурчик*
Компьютерная правка, оригинал-макет *В. М. Задоля*

Подписано в печать 23.09.2019. Формат 60×84 1/16. Бумага офсетная. Гарнитура «Таймс».
Отпечатано на ризографе. Усл. печ. л. 9,18. Уч.-изд. л. 10,0. Тираж 80 экз. Заказ 399.

Издатель и полиграфическое исполнение: учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники».
Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий №1/238 от 24.03.2014,
№2/113 от 07.04.2014, №3/615 от 07.04.2014.
Ул. П. Бровки, 6, 220013, г. Минск