

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ

**ХАРЬКОВСКИЙ НАЦИОНАЛЬНЫЙ ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ
ИМЕНИ СЕМЕНА КУЗНЕЦА**

**Методические рекомендации
к выполнению контрольных работ
по учебной дисциплине
"ЭКОНОМЕТРИКА"
для иностранных студентов
заочной формы обучения**

**Харьков
ХНЭУ им. С. Кузнеця
2016**

Утверждено на заседании кафедры высшей математики и экономико-математических методов.

Протокол № 4 от 30.11.2015 г.

Самостоятельное электронное текстовое сетевое издание

Составители: Э. Ю. Железнякова

И. Л. Лебедева

Л. А. Норик

Методические рекомендации к выполнению контрольных работ М 54 бот по учебной дисциплине "Эконометрика" для иностранных студентов заочной формы обучения : [Электронное издание] / сост. Э. Ю. Железнякова, И. Л. Лебедева, Л. А. Норик. – Харьков : ХНЭУ им. С. Кузнеця, 2016. – 75 с. (Рус. яз.)

Сборник содержит задачи к контрольным работам по учебной дисциплине и образцы выполнения нулевого варианта этих работ. Решение задач сопровождается кратким изложением теоретических положений, что помогает студентам заочной формы обучения получить теоретические знания и приобрести практические навыки их применения.

Рекомендовано для иностранных студентов заочной формы обучения.

Введение

Особенностью нынешнего этапа развития отечественной экономики является увеличение интереса специалистов к научному решению проблем с использованием экономико-математических методов и построения моделей на их основе. Экономико-математическое моделирование уже стало частью самой экономики и отражает сочетание знаний в области экономики, математики и кибернетики. Эконометрика занимает одно из ведущих мест в системе подготовки экономистов нового поколения, совместно с другими математическими и экономическими дисциплинами формирует новое экономическое мышление у будущих специалистов. Доскональное изучение учебной дисциплины "Эконометрика" имеет большое значение для общей подготовки специалистов по экономике, поскольку способствует созданию теоретической базы знаний и формированию практических навыков, которые необходимы для решения вопросов, связанных с исследованием экономических процессов и явлений, с целью оказания управляющего воздействия на их развитие.

В результате изучения нормативной дисциплины студенты приобретают компетентности в общей технологии моделирования и овладевают современными методами построения моделей объектов, явлений и процессов экономики. Они приобретают навыки реализации отдельных этапов моделирования, умение систематизировать, определять и оценивать значения исследуемых характеристик; усваивают базовые понятия, проблемы спецификации моделей и критерии эконометрического анализа; овладевают методами проверки выполнения основных предпосылок классической регрессионной модели.

Основной формой работы студента-заочника является самостоятельное рассмотрение теоретического материала и выполнение контрольных работ, что способствует лучшему усвоению программы. Предложенное учебное издание содержит задачи к контрольным работам и пример выполнения нулевого варианта.

Перед тем, как приступить к выполнению контрольной работы, необходимо тщательно изучить соответствующий теоретический материал по литературным источникам и методическим рекомендациям, список которых приведен в конце учебного издания. Помощь в решении каждой задачи студент может получить, рассмотрев примеры выполнения

нулевого варианта, которые приведены после перечня вариантов каждого задания контрольной работы.

Контрольная работа содержит четыре задания. Номер варианта студент определяет по номеру, согласно которому его фамилия записана в списке группы. Если работа частично или полностью не соответствует нужному варианту, то такая работа не засчитывается.

Формирование компетентностей современного специалиста предполагает овладение вычислительной техникой, поэтому вычисления, которые необходимо осуществлять в процессе построения эконометрических моделей, целесообразно выполнять на компьютере с применением программной среды *MS Excel*. Для проверки результатов вычисления можно применять встроенные функции и надстройки *MS Excel*, однако это не исключает необходимости применения алгоритма расчетов, предложенного в приведенном в данном пособии образце выполнения нулевого варианта каждого задания.

Контрольная работа оформляется в отдельной тетради. Решение каждого задания нужно начинать с новой страницы. Задачи необходимо располагать в порядке возрастания их номеров, сохраняя их номер по общему списку задач. Решению задания должно предшествовать его условие. Вычисления, выполненные на компьютере, рассматриваются только как приложения к основному решению задания. Они могут быть предоставлены в распечатанном виде.

Если контрольная работа не зачтена (это возможно, если она содержит ошибки, которые существенно влияют на результат), то ошибки необходимо исправить и результаты исправления привести в конце тетради или в новой тетради.

К экзамену по дисциплине могут быть допущены только те студенты, которые выполнили контрольную работу в полном объеме.

Задание 1. Эконометрическая модель парной регрессии

1.1. Задания для самостоятельного выполнения

По данным табл. 1.1 оценить параметры модели, которая характеризует зависимость цены товара Y (тыс. грн) от расходов на оплату труда X (грн). Построить линейную, степенную и экспоненциальную эконометрические модели, оценить их качество с помощью средней ошибки аппроксимации, а также по критерию Фишера. Сделать экономические выводы.

Таблица 1.1

Исходные данные для построения модели парной регрессии

Вариант 1		Вариант 2		Вариант 3		Вариант 4		Вариант 5	
Y	X	Y	X	Y	X	Y	X	Y	X
6,5	22,65	6,6	25,39	8,1	26,24	8,1	25,38	8,8	28,35
6,1	22,64	4,3	26,42	6,3	27,48	6,5	25,83	7,8	29,14
8,0	22,49	5,2	26,23	6,2	27,24	6,3	25,74	9,3	30,53
8,0	21,52	7,4	25,40	5,6	26,55	6,5	26,65	9,4	34,22
6,4	23,24	5,8	26,96	6,2	29,40	7,5	30,47	9,9	37,44
9,1	21,81	6,4	26,54	6,5	26,18	7,5	28,93	10,1	35,18
6,2	21,86	5,7	25,69	6,6	25,18	7,5	29,18	8,8	31,55
6,0	20,97	6,9	27,84	5,5	25,71	9,8	30,56	8,8	32,16
5,9	22,50	6,6	27,09	6,0	27,16	9,0	30,18	11,6	35,45
6,8	23,32	8,0	29,19	6,5	25,82	8,5	32,17	10,4	41,69
8,2	22,69	5,4	26,43	6,4	27,44	6,5	25,94	9,5	30,73
8,2	21,72	7,6	25,60	5,8	26,75	6,7	26,85	9,6	34,42
6,6	23,44	6,0	27,16	6,4	29,6	7,7	30,07	10,1	37,64
9,3	22,01	6,6	26,74	6,7	26,38	7,7	29,13	10,3	35,38
6,4	22,06	5,9	25,89	6,8	25,38	7,7	29,38	9,0	31,75
6,2	21,17	7,1	28,04	5,7	25,91	10,0	30,76	9,0	32,36
8,3	22,79	5,5	26,53	6,5	27,54	6,6	26,04	9,6	30,83
8,3	21,82	7,7	25,70	5,9	26,85	6,8	26,95	9,7	34,52
6,7	23,54	6,1	27,26	6,5	29,70	7,8	30,17	10,2	37,74
9,4	22,11	6,7	26,84	6,8	26,48	7,8	29,23	10,4	35,48
6,5	22,16	6,0	25,99	6,9	25,48	7,8	29,48	9,1	31,85
6,3	21,27	7,2	28,14	5,8	26,01	10,1	30,86	9,1	32,46
6,1	22,70	6,8	27,29	6,2	27,36	9,2	30,38	11,8	35,65
7,0	23,52	8,2	29,39	6,7	26,02	8,7	32,37	10,6	41,89
6,0	24,13	8,7	26,74	6,8	27,17	9,6	30,96	10,1	39,11
6,6	22,25	8,0	26,84	6,0	25,61	7,6	28,33	9,0	32,15

Продолжение табл. 1.1

Вариант 6		Вариант 7		Вариант 8		Вариант 9		Вариант 10	
Y	X	Y	X	Y	X	Y	X	Y	X
12,5	212,65	60,6	205,39	82,1	26,24	8,1	125,38	8,8	428,35
16,1	212,64	40,3	206,42	62,3	27,48	6,5	125,83	7,8	429,14
18,0	212,49	50,2	206,23	62,2	27,24	6,3	125,74	9,3	430,53
18,0	211,52	70,4	205,40	52,6	26,55	6,5	126,65	9,4	444,22
16,4	213,24	50,8	206,96	62,2	29,40	7,5	130,47	9,9	437,44
19,1	211,81	60,4	206,54	62,5	26,18	7,4	128,93	10,1	435,18
16,2	211,86	50,7	205,69	62,6	25,18	7,5	129,18	8,8	431,55
16,4	212,94	40,6	206,72	62,6	27,78	6,8	126,13	8,1	429,44
18,3	212,79	50,5	206,53	62,5	27,54	6,6	126,04	9,6	430,83
18,3	211,82	70,7	205,70	52,9	26,85	6,8	126,95	9,7	444,52
16,7	213,54	51,1	207,26	62,5	29,70	7,8	130,77	10,2	437,74
19,4	212,11	60,7	206,84	62,8	26,48	7,7	129,23	10,4	435,48
16,5	212,16	51,0	205,99	62,9	25,48	7,8	129,48	9,1	431,85
16,3	202,27	61,2	208,14	52,8	26,01	10,1	130,86	8,4	432,46
15,9	212,50	60,6	207,09	62,0	27,16	9,0	130,18	11,6	435,45
16,8	213,32	80,0	209,19	62,5	25,82	8,5	132,17	10,4	441,69
16,0	214,13	80,7	206,74	62,8	27,17	9,6	130,96	10,1	439,11
16,6	212,25	80,0	206,84	62,0	25,61	7,6	128,33	9,0	432,15

Вариант 11		Вариант 12		Вариант 13		Вариант 14		Вариант 15	
Y	X	Y	X	Y	X	Y	X	Y	X
26	21,20	35	25,39	45	26,24	50	25,38	60	28,35
28	22,64	36	26,42	46	27,48	52	25,83	66	29,14
25	22,49	39	26,23	46	27,24	51	25,74	65	30,53
28	21,52	40	25,40	49	26,55	63	26,65	70	34,22
30	23,24	35	26,96	40	29,40	60	30,47	74	37,44
32	21,81	30	26,54	42	26,18	62	28,93	72	35,18
35	21,86	35	25,69	40	25,18	65	29,18	70	31,55
30	20,97	40	27,84	41	25,71	68	30,56	75	32,16
31	25,64	39	29,42	49	30,48	55	28,83	69	32,14
28	25,49	42	29,23	49	30,24	54	28,74	68	33,53
31	24,52	43	28,40	52	29,55	66	29,65	73	37,22
33	26,24	38	29,96	43	32,40	63	33,47	77	40,44
35	24,81	33	29,54	45	29,18	65	31,93	75	38,18
38	24,86	38	28,69	43	28,18	68	32,18	73	34,55
33	23,97	43	30,84	44	28,71	71	33,56	78	35,16
42	22,50	42	27,09	42	27,16	62	30,18	72	35,45
45	23,32	45	29,19	45	25,82	65	32,17	76	41,69
50	24,13	43	26,74	50	27,17	60	30,96	70	39,11
35	22,25	39	26,84	43	25,61	55	28,33	65	32,15

Окончание табл. 1.1

Вариант 16		Вариант 17		Вариант 18		Вариант 19		Вариант 20	
Y	X	Y	X	Y	X	Y	X	Y	X
102,5	212,65	610,6	205,39	802,1	26,24	87,1	125,38	80,8	428,35
106,1	212,64	410,3	206,42	602,3	27,48	67,5	125,83	70,8	429,14
108,0	212,49	510,2	206,23	602,2	27,24	67,3	125,74	90,3	430,53
108,0	211,52	710,4	205,40	502,6	26,55	67,5	126,65	90,4	444,22
106,4	213,24	510,8	206,96	602,2	29,40	77,5	130,47	90,9	437,44
109,1	211,81	610,4	206,54	602,5	26,18	77,4	128,93	100,1	435,18
106,2	211,86	510,7	205,69	602,6	25,18	77,5	129,18	80,8	431,55
106,0	201,97	610,9	207,84	502,5	25,71	97,8	130,56	80,1	432,16
105,9	212,50	610,6	207,09	602,0	27,16	97,0	130,18	101,6	435,45
106,8	213,32	810,0	209,19	602,5	25,82	87,5	132,17	100,4	441,69
107,1	209,81	608,4	204,54	600,5	24,18	75,4	126,93	98,1	433,18
104,2	209,86	508,7	203,69	600,6	23,18	75,5	127,18	78,8	429,55
104,0	199,97	608,9	205,84	500,5	23,71	95,8	128,56	78,1	430,16
103,9	210,50	608,6	205,09	600,0	25,16	95,0	128,18	99,6	433,45
104,8	211,32	808,0	207,19	600,5	23,82	85,5	130,17	98,4	439,69
106,0	214,13	810,7	206,74	602,8	27,17	97,6	130,96	100,1	439,11
106,6	212,25	810,0	206,84	602,0	25,61	77,6	128,33	90,0	432,15

Вариант 21		Вариант 22		Вариант 23		Вариант 24		Вариант 25	
Y	X	Y	X	Y	X	Y	X	Y	X
206	22,12	35	215,39	45	206,24	50	25,38	60	218,35
208	22,64	36	216,42	46	207,48	52	25,83	66	219,14
205	22,49	39	216,23	46	207,24	51	25,74	65	310,53
208	21,52	40	215,40	49	206,55	63	26,65	70	314,22
300	23,24	35	216,96	40	209,40	60	30,47	74	317,44
302	21,81	30	216,54	42	206,18	62	28,93	72	315,18
305	21,86	35	215,69	40	205,18	65	29,18	70	311,55
300	20,97	40	217,84	41	205,71	68	30,56	75	312,16
402	22,50	42	217,09	42	207,16	62	30,18	72	315,45
305	28,24	40	221,96	45	214,40	65	35,47	79	322,44
307	26,81	35	221,54	47	211,18	67	33,93	77	320,18
310	26,86	40	220,69	45	210,18	70	34,18	75	316,55
305	25,97	45	222,84	46	210,71	73	35,56	80	317,16
407	27,50	47	222,09	47	212,16	67	35,18	77	320,45
410	28,32	50	224,19	50	210,82	70	37,17	81	416,69
500	24,13	43	216,74	50	207,17	60	30,96	70	319,11

1.2. Содержательная постановка задачи и ее математическая модель

Корреляция – вид статистической взаимосвязи между двумя или несколькими случайными величинами, при котором изменение одной из них приводит к изменению среднего значения другой. В зависимости от количества факторов, включенных в модель, различают модели парной и множественной корреляции. Описание корреляционной связи осуществляется с помощью уравнения регрессии. **Парная регрессия** – это уравнение $\hat{y}_x = f(x)$, которое описывает корреляционную зависимость между двумя факторами. В уравнении регрессии фактор x рассматривается как независимая, или **объясняющая**, переменная (**внешний**, или **экзогенный фактор**), а фактор y – зависимая переменная (**внутренний**, или **эндогенный фактор**), то есть **результативный признак**. Заметим, что обозначение \hat{y}_x отражает тот факт, что значение эндогенного фактора вычисляется по уравнению регрессии, а не получено в результате измерения.

По виду аппроксимирующей функции $f(x)$ различают модели линейной и нелинейной регрессии. Отметим, что модели нелинейной регрессии можно преобразовать в линейную, то есть **линеаризировать**, введя новые переменные. Уравнение линейной регрессии имеет вид:

$$\hat{y}_x = b_0 + b_1x, \quad (1.1)$$

где b_1 – **коэффициент регрессии**;

b_0 – свободный член.

Построение уравнения линейной регрессии сводится к оцениванию ее параметров по результатам исследования выборочной совокупности. Для этого используется **метод наименьших квадратов (МНК)**. МНК позволяет получить такие статистические оценки параметров уравнения регрессии, при которых сумма квадратов **ошибок модели**, то есть сумма квадратов отклонений эмпирических значений результативного признака от его значений, рассчитанных по уравнению регрессии, является минимальной. Условие минимизации функции ошибок модели приводит к решению **системы нормальных уравнений**, в результате чего получаем формулы для оценки параметров регрессионной модели.

Так, коэффициент регрессии определяется соотношением:

$$b_1 = \frac{\mu_{xy}}{\sigma_x^2} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2}, \quad (1.2)$$

где $\mu_{xy} = \overline{xy} - \bar{x} \cdot \bar{y} = \frac{\sum_{i=1}^n x_i y_i}{n} - \bar{x} \cdot \bar{y}$ – **коэффициент ковариации**;

$\sigma_x^2 = \overline{x^2} - \bar{x}^2 = \frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2$ – дисперсия фактора X в генеральной совокупности;

$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ – выборочная средняя экзогенного фактора X ;

$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$ – выборочная средняя эндогенного фактора Y ;

n – объем выборочной совокупности.

Свободный член уравнения регрессии определяется соотношением:

$$b_0 = \bar{y} - b_1 \bar{x}. \quad (1.3)$$

В рамках линейной модели для оценки тесноты линейной корреляционной связи используют **коэффициент корреляции Пирсона**:

$$r_{xy} = \frac{\mu_{xy}}{\sigma_x \sigma_y} = b_1 \frac{\sigma_x}{\sigma_y}. \quad (1.4)$$

Значение коэффициента корреляции может изменяться в пределах $-1 \leq r_{xy} \leq 1$. Связь считается тесной, если $|r_{xy}| \geq 0,7$. Если $|r_{xy}| \leq 0,35$, то связь считается статистически незначимой, и ею можно пренебречь. Положительное значение r_{xy} означает, что с увеличением одного из факторов среднее значение другого фактора увеличивается. Если r_{xy} отрицательно, то с увеличением одного из факторов среднее значение другого фактора уменьшается.

Для того чтобы определить качество регрессионной модели независимо от формы связи, рассчитывают долю изменчивости эндогенного фактора в общей изменчивости этого фактора, которую в рамках построенной модели можно объяснить изменением экзогенного фактора. Для этого вычисляют **коэффициент детерминации**:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}, \quad (1.5)$$

а также **индекс корреляции**:

$$\eta_{xy} = \sqrt{1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}}, \quad (1.6)$$

где $\frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$ – доля суммы квадратов остатков модели $\sum (y_i - \hat{y}_i)^2$ в общей сумме квадратов изменчивости эндогенного фактора $\sum (y_i - \bar{y})^2$.

Для линейной модели имеет место соотношение: $R^2 = r_{xy}^2$.

Кроме того, для оценки качества построенной модели рассчитывают **среднюю ошибку аппроксимации**, которая является средним арифметическим относительных отклонений расчетных значений эндогенного фактора от его эмпирических значений по каждому наблюдению:

$$\bar{A} = \frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100 \%}{n}, \quad (1.7)$$

где y_i – эмпирическое значение эндогенного фактора в i -м наблюдении;
 \hat{y}_i – теоретическое значение эндогенного фактора, рассчитанное по уравнению регрессии при подстановке в него значения x_i .

Модель считается качественной, если средняя ошибка аппроксимации не превышает 8 – 10 %.

Для того чтобы узнать, на сколько процентов в среднем изменится значение эндогенного фактора относительно своей выборочной средней

при условии, что экзогенный фактор изменится на 1% от своей выборочной средней, вычисляются **коэффициент эластичности**:

$$\varepsilon = f'(x) \frac{\bar{x}}{y} = b_1 \frac{\bar{x}}{y}. \quad (1.8)$$

Проверка значимости тесноты корреляционной связи осуществляется по критерию **Фишера – Снедекора** (F -критерий). Рассматривается основная гипотеза $H_0: r_{xy} = 0$, то есть предполагается, что связь между факторами, включенными в модель, является статистически незначимой, При этом альтернативной является гипотеза $H_1: r_{xy} \neq 0$, в соответствии с которой связь является значимой. Для этого для выбранного **уровня значимости** ($\alpha = 0,05$ или $\alpha = 0,01$) сравнивают табличное значение критерия Фишера $F_{табл}$ и его фактическое

значение, которое вычисляется по формуле: $F_{факт} = \frac{r_{xy}^2}{1 - r_{xy}^2} (n - 2)$. По

справочным таблицам (табл. А.1 приложения А, табл. Б.1 приложения Б) находят значения статистик критерия Фишера: $F_{0,01}(1; n - 2)$ и $F_{0,05}(1; n - 2)$. Если $F_{факт} > F_{0,01}$, то основная гипотеза отвергается в пользу альтернативной, и связь между факторами следует считать статистически значимой. Если $F_{факт} < F_{0,05}$, то основную гипотезу нет оснований отвергнуть, и связь считается статистически незначимой. Если $F_{0,05} < F_{факт} < F_{0,01}$, то это область неопределенности, и нужны дополнительные исследования.

Статистическая значимость каждого из параметров уравнения регрессии отдельно, а также коэффициента корреляции проверяются по **критерию Стьюдента** (t -критерий), который для каждой из этих характеристик рассчитывается из соотношения:

$$t_r = \frac{r}{m_r}, \quad t_{b_j} = \frac{b_j}{m_{b_j}} \quad \text{при } j \in \overline{0, 1}, \quad (1.9)$$

где m_{b_j} – средние ошибки параметров линейной регрессии;

m_r – средняя ошибка коэффициента корреляции.

Средние ошибки соответствующих величин определяют по формулам:

$$m_{b_0} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2} \cdot \frac{\sum x_i^2}{n \sum (x_i - \bar{x})^2}};$$

$$m_{b_1} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2} \cdot \frac{1}{\sum (x_i - \bar{x})^2}}; \quad (1.10)$$

$$m_r = \sqrt{\frac{1 - r_{xy}^2}{n-2}}.$$

Сравнивая фактическое значение критерия Стьюдента и его табличное значение (табл. В.1 приложения В) при количестве степеней свободы $k = n - 2$, мы принимаем или отвергаем основную гипотезу H_0 о незначимости рассматриваемой характеристики. Если $t_{табл} < t_{факт}$, то гипотеза H_0 отклоняется, то есть исследуемая характеристика является статистически значимой, а если $t_{табл} > t_{факт}$, то признается статистическая значимость исследуемой характеристики.

1.3. Пример выполнения типового задания № 1

Исходные данные. В табл. 1.2 приведены эмпирические данные о расходах на оплату труда x_i (экзогенный фактор) и цене продукции y_i (эндогенный фактор), полученные по результатам семи измерений.

Таблица 1.2

Исходные данные задачи

№ измерения	X	Y
1	45,1	68,8
2	47,2	54,3
3	55,2	49,3
4	57,2	59,9
5	58,8	55
6	59,0	61,2
7	61,8	56,7

По эмпирическим данным построим график, отражающий зависимость между внешним и внутренним факторам (рис. 1.1).

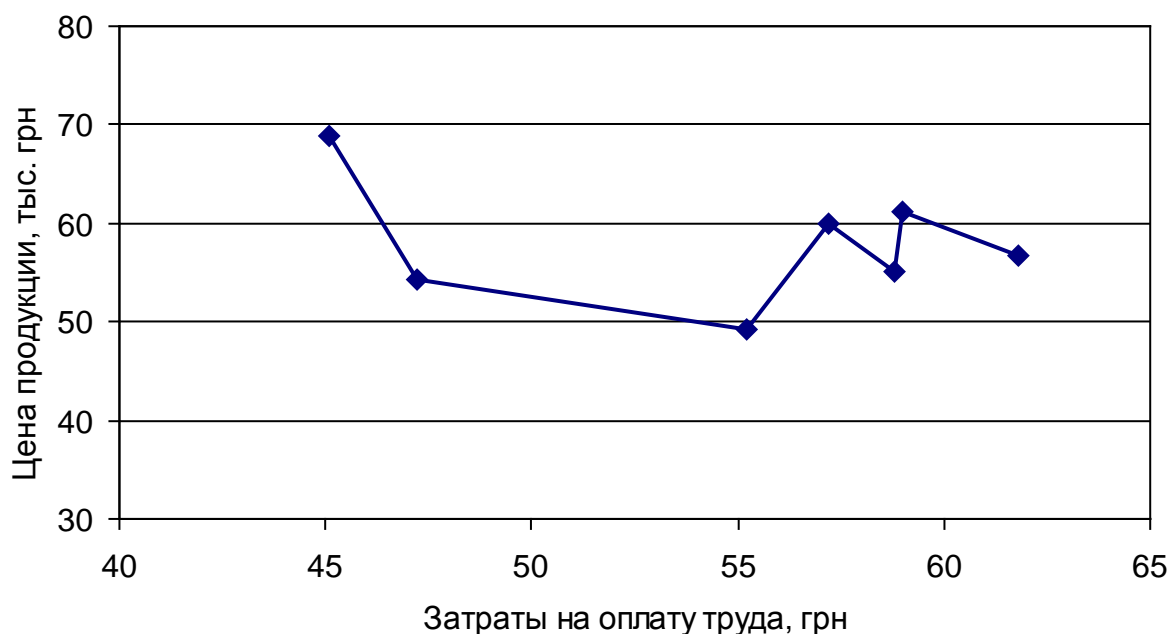


Рис. 1.1. Влияние оплаты труда на цену продукции

Анализ вида эмпирической зависимости свидетельствует, что при построении эконометрической модели можно сделать несколько предположений относительно вида функции, с помощью которой описывается эта зависимость. Это может быть как линейная, так и степенная или показательная функции. Построим три различные эконометрические модели, принимая в качестве функции, с помощью которой осуществляется аппроксимация, одну из этих функций и проведем оценивание качества каждой модели по средней ошибке аппроксимации и критерию Фишера, а также определим статистическую значимость параметров модели.

Рассмотрим сначала линейную аппроксимацию, поскольку другие модели с помощью преобразований и использования вспомогательных переменных сводятся к построению линейной модели. В предположении линейной аппроксимации уравнение регрессии имеет вид (1.1):

$$\hat{y}_x = b_0 + b_1x.$$

По эмпирическим данным выполним вспомогательные вычисления, результаты которых приведены в табл. 1.3.

**Вспомогательные вычисления для оценки параметров
линейной модели "Цена продукции – затраты на оплату труда"**

№	x_i	y_i	$x_i y_i$	x_i^2	y_i^2	\hat{y}_i	$\left \frac{y_i - \hat{y}_i}{y_i} \right \cdot 100\%$
1	45,1	68,8	3 102,88	2 034,01	4 733,44	61,3	10,9
2	47,2	54,3	2 562,96	2 227,84	2 948,49	60,5	11,5
3	55,2	49,3	2 721,36	3 047,04	2 430,49	57,8	17,2
4	57,2	59,9	3 426,28	3 271,84	3 588,01	57,1	4,7
5	58,8	55	3 234,00	3 457,44	3 025,00	56,5	2,8
6	59,0	61,2	3 610,80	3 481,00	3 745,44	56,5	7,7
7	61,8	56,7	3 504,06	3 819,24	3 214,89	55,5	2,1
Σ	384,3	405,2	22 162,34	21 338,41	23 685,76	405,2	57,0

Последняя строка табл. 1.3 содержит суммы по каждому из столбцов и используется для вычисления соответствующих средних значений. По данным полученной расчетной таблицы вычисляем выборочные средние, ковариацию, дисперсию и параметры уравнения модели линейной регрессии:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{384,3}{7} = 54,9; \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{405,2}{7} = 57,89;$$

$$\mu_{xy} = \overline{xy} - \bar{x} \cdot \bar{y} = \frac{\sum_{i=1}^n x_i y_i}{n} - \bar{x} \cdot \bar{y} = \frac{22162,34}{7} - 54,9 \cdot 57,89 = -11,88;$$

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2 = \frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2 = \frac{21338,41}{7} - (54,9)^2 = 34,33;$$

$$b_1 = \frac{\mu_{xy}}{\sigma_x^2} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} = -0,35; \quad b_0 = \bar{y} - b_1 \bar{x} = 76,88.$$

Таким образом, получаем уравнение линейной регрессии:

$$\hat{y}_x = 76,88 - 0,35x.$$

Рассмотрим, какую информацию можно получить из значения коэффициента регрессии. Так как $b_1 < 0$, то увеличение фактора X приводит к уменьшению значения фактора Y . То есть корреляционная связь между исследуемыми факторами является обратной. По абсолютной величине коэффициента регрессии можно сделать вывод, что увеличение затрат на оплату труда на 1 единицу приводит к уменьшению цены продукции на 0,35 единиц.

Для оценки тесноты корреляционной связи вычислим коэффициент корреляции:

$$r_{xy} = \frac{\mu_{xy}}{\sigma_x \sigma_y} = b_1 \frac{\sigma_x}{\sigma_y} = -0,357.$$

Так как $|r| > 0,35$, то связь можно считать значимой.

Вычислим коэффициент детерминации. В случае линейной регрессии имеем: $R^2 = r_{xy}^2 = 0,127$. Это означает, что изменчивость эндогенного фактора Y только на 12,7 % можно объяснить изменчивостью экзогенного фактора X .

Для проверки качества выбранной модели рассчитаем среднюю ошибку аппроксимации:

$$\bar{A} = \frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100 \%}{n} = 8,1 \%,$$

то есть в среднем расчетные значения фактора Y отклоняются от его фактических значений на 8,1%, что является вполне допустимым.

Теперь построим эконометрическую модель для случая, когда функция, с помощью которой осуществляется аппроксимация, является показательной. В этом случае уравнение регрессии имеет вид:

$$\hat{y}_x = b_0 \cdot b_1^x.$$

Для оценки параметров уравнения регрессии проведем линеаризацию исходного уравнения путем его логарифмирования:

$$\ln \hat{y} = \ln(b_0 \cdot b_1^x) \Rightarrow \ln \hat{y} = \ln b_0 + x \cdot \ln b_1.$$

Если ввести вспомогательную переменную $\hat{Y} = \ln \hat{y}$ и обозначить $B_0 = \ln b_0$, $B_1 = \ln b_1$, то относительно них получим линейное уравнение:

$$\hat{Y} = B_0 + B_1 \cdot x.$$

Итак, мы осуществили преобразования показательной модели в линейную. Рассчитываем параметры модели, которую получили в результате линеаризации показательной функции, по методу МНК:

$$B_1 = \frac{\mu_{xY}}{\sigma_x^2} = \frac{\overline{xY} - \bar{x} \cdot \bar{Y}}{\overline{x^2} - \bar{x}^2}; \quad B_0 = \bar{Y} - B_1 \cdot \bar{x}.$$

Выполненные вычисления приведены в табл. 1.4.

Таблица 1.4

Вспомогательные вычисления оценки параметров показательной модели "Цена продукции – затраты на оплату труда"

№	x	y	$Y = \ln y$	xY	x^2	Y^2	\hat{Y}	\hat{y}	$\left \frac{y - \hat{y}}{y} \right \cdot 100\%$
1	45,1	68,8	4,23	190,83	2 034,01	17,90	4,12	61,84	10,12
2	47,2	54,3	3,99	188,54	2 227,84	15,96	4,11	61,19	12,69
3	55,2	49,3	3,90	215,17	3 047,04	15,19	4,07	58,79	19,25
4	57,2	59,9	4,09	234,10	3 271,84	16,75	4,06	58,21	2,83
5	58,8	55,0	4,01	235,63	3 457,44	16,06	4,06	57,74	4,99
6	59,0	61,2	4,11	242,73	3 481,00	16,93	4,06	57,69	5,74
7	61,8	56,7	4,04	249,53	3 819,24	16,3	4,04	56,88	0,32
Σ	384,3	405,2	28,38	1 556,54	21 338,41	115,09			55,94

Получаем значения параметров модели:

$$B_1 = \frac{\mu_{xY}}{\sigma_x^2} = \frac{\overline{xY} - \bar{x} \cdot \bar{Y}}{\overline{x^2} - \bar{x}^2} = -0,005; \quad B_0 = \bar{Y} - B_1 \cdot \bar{x} = 4,35.$$

Уравнение регрессии в новых переменных имеет вид:

$$\hat{Y} = 4,35 - 0,005 \cdot x.$$

Вернемся к исходным переменным. Для этого выполним потенцирование полученного уравнения. Итак, при использовании в качестве аппроксимирующей показательную функцию эконометрическая модель имеет вид:

$$\hat{y} = e^{4,35} e^{-0,005x} = 77,5 \cdot 0,995^x.$$

Для определения качества показательной модели рассчитаем среднюю ошибку аппроксимации:

$$\bar{A} = \frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100 \%}{n} = 7,99 \%,$$

то есть в среднем расчетные значения отклоняются от эмпирических на 7,98 %, соответственно, показательная модель лучше описывает эмпирические данные, чем линейная.

Рассмотрим еще один случай. Предположим, что для аппроксимации была выбрана степенная функция. Получаем уравнение регрессии:

$$\hat{y} = b_0 \cdot x^{b_1}.$$

Преобразуем его к линейному виду путем логарифмирования:

$$\ln \hat{y} = \ln(b_0 \cdot x^{b_1}).$$

Вводим новые обозначения и получаем:

$$\hat{Y} = B_0 + B_1 \cdot X,$$

где $\hat{Y} = \ln \hat{y}$, $B_0 = \ln b_0$, $B_1 = b_1$, $X = \ln x$.

Для вычисления параметров модели, которую получили в результате линеаризации степенной функции, применяем соотношение:

$$B_1 = \frac{\mu_{XY}}{\sigma_X^2} = \frac{\overline{XY} - \bar{X} \cdot \bar{Y}}{\overline{X^2} - \bar{X}^2}; \quad B_0 = \bar{Y} - B_1 \cdot \bar{X}.$$

Результаты вспомогательных вычислений, необходимые для расчета параметров модели, приведены в табл. 1.5.

Таблица 1.5

Вспомогательные вычисления для оценки параметров степенной модели "Цена продукции – затраты на оплату труда"

№	x	y	$Y =$ $= \ln y$	$X =$ $= \ln x$	$X \cdot Y$	X^2	\hat{Y}	\hat{y}	$\left \frac{y - \hat{y}}{y} \right \cdot 100 \%$
1	45,1	68,8	4,23	3,81	16,12	14,52	4,24	60,59	11,94
2	47,2	54,3	3,99	3,85	15,36	14,82	4,20	59,85	10,21
3	55,2	49,3	3,90	4,01	15,64	16,08	4,04	57,37	16,37
4	57,2	59,9	4,09	4,05	16,56	16,40	4,00	56,82	5,14
5	58,8	55,0	4,01	4,07	16,32	16,56	3,98	56,40	2,54
6	59,0	61,2	4,11	4,08	16,77	16,65	3,97	56,35	7,93
7	61,8	56,7	4,04	4,12	16,64	16,97	3,93	55,65	1,86
Σ	384,3	405,2	28,37	27,99	113,42	112,01			55,99

$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n} = 4,05; \quad \bar{X} = \frac{\sum_{i=1}^n X_i}{n} = 4,0;$$

$$\mu_{XY} = \frac{\sum_{i=1}^n X_i \cdot Y_i}{n} - \bar{X} \bar{Y} = -0,0034; \quad \sigma_X^2 = \frac{\sum_{i=1}^n X_i^2}{n} - \bar{X}^2 = 0,0124.$$

Отсюда:

$$B_1 = -0,27; \quad B_0 = 5,13.$$

Итак, уравнение регрессии с использованием вспомогательных переменных имеет вид:

$$\hat{Y} = 5,13 - 0,27 \cdot X.$$

Для возврата к исходным переменным выполним потенцирование и запишем уравнение регрессии в виде степенной функции:

$$\hat{y} = e^{5,13} \cdot x^{-0,27} = 169,44 \cdot x^{-0,27}.$$

Для определения качества выбранной модели рассчитаем среднюю ошибку аппроксимации. Она равна:

$$\bar{A} = \frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100 \%}{n} = 8,0 \%,$$

то есть в среднем расчетные значения отклоняются от фактических на 8,0 %, что является вполне допустимым.

Итак, по средней ошибкой аппроксимации все три модели показали почти одинаковые результаты.

Вернемся к анализу линейной модели. С целью исследования статистической значимости уравнения линейной модели в целом вычислим эмпирическое значение критерия Фишера – Снедекора:

$$F_{\text{факт}} = \frac{r_{xy}^2}{1 - r_{xy}^2} (n - 2) = \frac{0,127}{1 - 0,127} (7 - 2) = 0,73.$$

По справочным таблицам находим критические значения F -критерия при различных уровнях значимости. Так, $F_{0,01}(1; n - 2) = F_{0,01}(1; 5) = 16,3$ (см. табл. А.1 приложения А) и $F_{0,05}(1; n - 2) = F_{0,05}(1; 5) = 6,61$ (см. табл. Б.1 приложения Б). Поскольку $F_{\text{факт}} < F_{0,05}$, то основную гипотезу H_0 об отсутствии статистически значимой корреляционной связи нет оснований отбросить. Иными словами, изменчивость эндогенного фактора нельзя объяснить влиянием того экзогенного фактора, который предполагается в соответствии со спецификацией данной модели. Эта изменчивость связана с воздействием факторов, не учтенных в этой модели.

Для проверки значимости именно корреляционной связи применим статистику Стьюдента. По основной гипотезе $H_0: r_{xy} = 0$. Для проверки этой гипотезы при альтернативной гипотезе $H_1: r_{xy} \neq 0$ определим случайную ошибку коэффициента корреляции. Она равна:

$$m_r = \sqrt{\frac{1 - r_{xy}^2}{n - 2}} = \sqrt{\frac{1 - 0,127}{5}} = 0,42.$$

Теперь вычисляем эмпирическое значение критерия Стьюдента для проверки значимости коэффициента корреляции:

$$t_r = \frac{|r|}{m_r} = \frac{0,36}{0,42} = 0,86.$$

По справочной таблице (см. табл. В.1 приложения В) определяем статистику Стьюдента для количества степеней свободы $k = n - 2 = 7 - 2 = 5$ при уровне значимости $\alpha = 0,05$. Получаем: $t_{0,05}(5) = 2,57$. Поскольку имеет место соотношение $t_{\text{факт}} < t_{\text{табл}}$, то гипотезу H_0 нет причин отбросить. Итак, корреляционная связь между факторами, которые рассматриваются в соответствии со спецификацией модели, является статистически незначимой. Таким образом, выводы о том, что зависимость стоимости продукции от расходов на оплату труда является статистически незначимой, полученные в результате проверки с использованием статистических критериев Фишера – Снедекора и Стьюдента, совпадают.

Проверим с помощью t -критерия статистическую значимость каждого из параметров линейной модели. Для каждого параметра модели проверке подлежит основная гипотеза H_0 , в соответствии с которой $b_1 = 0$ ($b_0 = 0$). По табл. 1.3 вычислим сумму квадратов остатков модели:

$$\begin{aligned} \sum (y_i - \hat{y}_i)^2 &= (68,8 - 61,3)^2 + (61,2 - 56,5)^2 + (59,9 - 57,1)^2 + \\ &+ (56,7 - 55,5)^2 + (55,0 - 56,5)^2 + (54,3 - 60,5)^2 + (49,3 - 57,8)^2 = 200,55. \end{aligned}$$

Теперь определяем случайные ошибки для каждого из параметров модели. Получаем:

$$m_{b_1} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2} \cdot \frac{1}{\sum (x_i - \bar{x})^2}} = \sqrt{\frac{200,55}{5} \cdot \frac{1}{7 \cdot 34,33}} = 0,41;$$

$$m_{b_0} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2} \cdot \frac{\sum x_i^2}{n \sum (x_i - \bar{x})^2}} = \sqrt{\frac{200,55}{5} \cdot \frac{21338,41}{7 \cdot 34,33}} = 22,56.$$

По этим данным вычисляем эмпирические значения критерия Стьюдента для каждого из параметров модели.

Получаем:

$$t_{b_1} = \frac{|b_1|}{m_{b_1}} = \frac{0,35}{0,41} = 0,85; \quad t_{b_0} = \frac{b_0}{m_{b_0}} = \frac{76,88}{22,56} = 3,41.$$

По табл. В.1 приложения В находим, что табличное значение статистики Стьюдента для числа степеней свободы $k = n - 2 = 7 - 2 = 5$ и уровня значимости $\alpha = 0,05$ равно $t_{0,05}(5) = 2,57$.

Проверим основную гипотезу $H_0: b_1 = 0$, если альтернативной является $H_1: b_1 \neq 0$. Заметим, что именно значимость коэффициента регрессии b_1 определяет тесноту корреляционной связи между факторами, которые учитываются в модели в соответствии с ее спецификацией. Поскольку при проверке значимости параметра b_1 выполняется неравенство $t_{\text{факт}} < t_{\text{табл}}$, то гипотезу H_0 нет оснований отклонить. Это означает, что коэффициент регрессии b_1 является статистически незначимым, то есть влияние экзогенного фактора на эндогенный следует считать несущественным. Это полностью совпадает с предыдущими выводами.

Проверим гипотезу $H_0: b_0 = 0$ при альтернативной $H_1: b_0 \neq 0$. Так как $t_{\text{факт}} > t_{\text{табл}}$, то эту гипотезу отвергаем в пользу альтернативной, то есть параметр b_0 значим. Итак, имеем линейную модель в виде:

$$\hat{y} = \bar{y},$$

то есть эндогенный фактор сохраняется на уровне выборочной средней и не зависит от экзогенного фактора, который учитывается в данной модели.

Вывод: по величине средней ошибки аппроксимации среди представленных эконометрических моделей парной регрессии, описывающих влияние затрат на оплату труда на цену продукции, более качественной является эконометрическая модель, согласно которой функция, используемая для аппроксимации, является линейной. Однако проверка значимости коэффициента корреляции, уравнения в целом, а также его параметров показала, что независимо от затрат на оплату труда цена продукции остается на уровне выборочной средней. Эти данные можно использовать для краткосрочного прогноза цены продукции.

Задание 2. Эконометрическая модель многофакторной регрессии

2.1. Задания для самостоятельного выполнения

По эмпирическим данным о производительности труда Y (тыс. грн), фондообеспеченности X_1 (тыс. грн) и стаже работы X_2 (года), которые представлены в табл. 2.1, необходимо:

построить эконометрическую модель многофакторной регрессии в стандартизированных и натуральных переменных;

проверить значимость регрессии по критерию Фишера;

установить целесообразность присутствия каждого из факторов в уравнении регрессии с помощью частных критериев Фишера;

сделать выводы.

Таблица 2.1

Исходные данные для построения модели многофакторной регрессии

№ варианта	Показатель	Значения показателей (в соответствующих единицах)														
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1	Y	36,1	19,9	19,4	20,2	19,2	20,5	20,2	25,1	20,5	19,4	20,6	20,8	18,6	18,1	
	X_1	98	260	318	302	231	200	200	140	225	232	231	200	225	258	
	X_2	14,4	15,5	13,2	12,8	19,2	18,2	15,8	15,4	17,2	17,2	15,8	16,7	18,7	15,1	
2	Y	66	110	140	139	105	95	85	88	100	90	105	85	110	120	
	X_1	110	140	139	145	139	140	130	129	138	135	155	142	125	150	
	X_2	19,9	19,4	20,2	19,2	18,1	17,5	17	17,6	16,5	18,2	16,9	15,5	19,2	18,5	
3	Y	260	318	302	305	302	318	305	302	351	318	350	351	267	360	
	X_1	365	375	350	345	305	400	384	372	395	383	430	405	360	390	
	X_2	19,2	15,4	13,2	16,2	18	17,2	16,7	13,2	14,9	15,7	12,9	16,4	14,9	14,7	
4	Y	231	231	151	112	151	225	156	105	105	135	135	112	105	105	
	X_1	295	315	285	260	240	235	275	252	273	286	336	215	221	258	
	X_2	10,9	11,0	4,5	8,7	6,8	4,8	5,9	12,4	9,1	9,5	13,5	6,3	6,6	15,0	

Продолжение табл. 2.1

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
5	Y	231	231	151	112	151	225	156	105	105	135	135	86	105	105
	X ₁	105	115	90	85	90	100	105	70	70	84	84	64	63	63
	X ₂	35,5	34,5	25,6	27,5	26,3	34,3	27,5	20,2	19,2	14,5	23,8	17,2	15,3	12,5
6	Y	20	29	25	20	21	23	27	29	27	21	35	28	30	21
	X ₁	107	124	127	105	182	182	125	132	128	164	154	122	126	108
	X ₂	7,2	9,0	9,7	12,7	14,2	16,0	6,2	8,1	9,0	17,1	6,6	11,6	7,4	12,5
7	Y	334	224	278	329	308	309	239	228	225	317	175	244	224	274
	X ₁	161	186	205	165	130	135	197	180	169	155	244	150	188	200
	X ₂	31	18	20	24	34	15	27	25	24	32	17	19	12	29
8	Y	164	122	99	232	174	122	152	174	160	274	178	183	157	115
	X ₁	217	929	876	223	146	200	115	142	395	301	259	136	200	851
	X ₂	31,4	24,3	22,5	37,0	30,9	24,0	26,0	31,7	25,8	34,2	32,0	20,7	18,2	22,3
9	Y	164	122	99	232	174	123	152	174	122	274	178	183	157	115
	X ₁	590	776	929	452	141	194	121	979	143	601	140	239	189	109
	X ₂	22	23	25	24	23	20	22	29	20	35	23	23	24	21
10	Y	147	98	121	144	182	182	98	91	110	74	147	182	98	98
	X ₁	106	90	97	107	142	160	90	81	90	66	140	165	115	90
	X ₂	34	24	27	25	18	16	23	20	15	17	27	19	23	21
11	Y	43,1	20,3	17	21,6	16,2	31,5	31,9	25,4	27,2	37,3	41,5	34,3	44,3	43,4
	X ₁	90	131	163	121	163	89	89	183	141	91	98	97	90	90
	X ₂	21,5	15,9	13,6	15,7	15,8	14,9	14	20,1	24,8	14,7	14,7	15,8	21,7	23,7
12	Y	164	122	99	232	174	123	152	174	122	274	178	183	157	115
	X ₁	590	779	924	427	171	192	161	969	134	601	140	239	189	100
	X ₂	23	25	21	24	28	18	22	29	20	37	23	25	24	27
13	Y	48	103	125	115	133	71	71	77	71	69	76	78	48	48
	X ₁	135	133	76	103	77	67	48	115	71	72	73	69	76	78
	X ₂	17,2	16,2	25,1	20,3	25,4	36,4	43,1	21,6	31,5	31,9	35,2	37,3	33,3	34,3
14	Y	163	164	145	131	183	121	87	121	89	88	141	91	98	97
	X ₁	310	320	360	280	353	295	185	280	190	195	319	200	214	211
	X ₂	13,6	11,1	21,5	14,9	14	14,7	21,7	23,7	15,3	14,2	15,3	24,6	15,1	15,7
15	Y	163	145	90	109	102	98	108	94	109	105	95	97	126	122
	X ₁	190	230	316	200	215	310	210	305	214	238	308	332	250	200
	X ₂	14,2	25,2	6,6	17,4	7,2	8,3	7,5	10,2	17,6	6,9	8,4	17,0	7,2	21,3

Окончание табл. 2.1

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
16	Y	163	145	90	109	108	111	90	91	101	105	106	97	122	121
	X ₁	125	76	48	71	71	76	48	48	62	74	74	52	88	115
	X ₂	34	18	20	30	24	14	25	35	21	38	19	21	25	30
17	Y	6,3	5,8	5,9	5,9	6,1	6,2	5,7	5,8	5,9	6,8	6,5	5,9	6,2	5,7
	X ₁	3,3	2,7	3,1	2,9	3,1	3,2	2,5	2,9	2,5	3,6	3,2	2,7	3,1	2,5
	X ₂	15,9	25,1	5,2	8,1	10,8	12,1	16,3	6,2	4,8	6,1	10,7	9,3	5,2	16,3
18	Y	2,5	1,9	2,1	1,8	2,2	2,1	1,7	1,8	1,8	2,5	2	1,9	2,1	2
	X ₁	3,2	3,1	3,1	2,9	3,2	3,1	2,5	2,9	2,5	2,9	2,9	2,8	3,1	2,9
	X ₂	24,7	14,5	8,9	4,2	10,6	6,5	9,1	12,2	17,3	16,9	9,8	4,3	7,7	11,7
19	Y	1,4	1,5	1,5	1,3	1,5	1,3	1,6	0,9	1,3	1,4	0,9	1,5	1,1	1,4
	X ₁	7,1	6,4	6,9	5,5	6,5	5,7	6,3	4,9	6,6	5,2	4,9	5,9	5,9	6,1
	X ₂	11	9	16	25	34	16	20	30	24	14	25	35	45	34
20	Y	4	4,6	4,7	3,9	4,3	4,4	4,1	4,2	4,4	5,1	5,9	5,6	4,5	5,5
	X ₁	1,3	1,5	1,4	1,1	1,3	1,3	1,3	1,3	1,2	1,9	2,1	1,8	1,7	1,8
	X ₂	12	5	11	18	22	29	20	17	23	25	24	19	12	8
21	Y	106	90	97	107	142	160	90	81	90	161	66	96	74	125
	X ₁	374	243	225	331	394	240	262	317	278	342	320	287	232	220
	X ₂	24	14	25	35	45	38	19	21	14	15	18	10	8	19
22	Y	20	22	25	20	21	23	27	19	27	15	13	23	28	29
	X ₁	164	122	111	102	160	122	132	170	125	104	128	101	175	115
	X ₂	10,8	12,1	16,3	6,2	4,8	6,1	10,7	9,3	5,2	10,1	18,3	13,6	20,4	9,5
23	Y	106	101	107	158	137	101	107	110	104	168	140	123	115	109
	X ₁	147	98	121	144	182	182	98	91	110	74	147	182	98	98
	X ₂	24	14	25	35	21	38	19	21	19	16	14	15	10	9
24	Y	20	29	25	20	21	23	27	29	27	35	23	23	28	29
	X ₁	147	98	121	144	182	182	98	91	110	74	147	182	98	98
	X ₂	4,2	10,6	6,5	9,1	12,2	17,3	16,9	9,8	4,3	6,6	14,0	16,5	11,5	9,0
25	Y	147	98	121	144	182	182	98	91	110	74	147	182	98	98
	X ₁	106	90	97	107	142	160	90	81	90	66	140	165	115	90
	X ₂	14,0	10,2	17,0	30,5	19,2	16,2	25,1	20,3	25,4	16,4	23,1	21,6	28,5	31,9

2.2. Содержательная постановка задачи и ее математическая модель

Если рассматривать совокупность X_1 , X_2 и Y как трехмерную случайную величину, то она имеет девять основных числовых характеристик, а именно: выборочные средние каждого из трех компонентов, их средние квадратические отклонения и парные коэффициенты корреляции между всеми компонентами случайной величины. Приведем формулы для вычисления основных числовых характеристик эмпирического распределения такой случайной величины. Так, для каждого из компонентов выборочная средняя определяется соотношением:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}, \quad (2.1)$$

где x_i – значение i -й варианты ($i = \overline{1, n}$) компонента X_j (в данном случае это любой из компонентов X_1 , X_2 или Y) трехмерной случайной величины;

n – объем выборочной совокупности (количество вариантов).

Выборочная средняя является несмещенной оценкой средней генеральной совокупности.

Для вычисления дисперсии любого из компонентов трехмерной случайной величин по несгруппированным данным применяют формулу:

$$\sigma_x^2 = \overline{x^2} - (\bar{x})^2 = \frac{\sum_{i=1}^n x_i^2}{n} - (\bar{x})^2. \quad (2.2)$$

Отсюда получаем среднее квадратическое отклонение:

$$\sigma_x = \sqrt{\sigma_x^2}. \quad (2.3)$$

Поскольку по формуле (2.2) получаем смещенную оценку дисперсии генеральной совокупности, то необходимо ввести поправку на смещение. Соответственно, получаем исправленную дисперсию:

$$S_x^2 = \frac{n}{n-1} \cdot \sigma_x^2. \quad (2.4)$$

Исправленное среднее квадратическое отклонение составляет:

$$S_x = \sqrt{S_x^2}. \quad (2.5)$$

Выборочная средняя и среднее квадратическое отклонение характеризуют каждый из компонентов трехмерной случайной величины отдельно. Характеристиками корреляционной связи между компонентами этой случайной величины является парные коэффициенты корреляции. Так, парный коэффициент корреляции между компонентами X и Y вычисляется по формуле:

$$r_{xy} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y}, \quad (2.6)$$

где \overline{xy} – среднее произведения случайных величин.

Эконометрическую модель может построить как в натуральных, так и в стандартизированных переменных. **Стандартизованная переменная**, которая соответствует натуральной переменной X_j (это любой из компонентов X_1 , X_2 или Y трехмерной случайной величины), задается соотношением:

$$t_{xj} = \frac{X_j - \bar{x}_j}{S_{xj}}. \quad (2.7)$$

В стандартизированных переменных эконометрическая модель имеет вид:

$$\hat{t}_y = \beta_1 \cdot t_{x_1} + \beta_2 \cdot t_{x_2}, \quad (2.8)$$

где β_1 и β_2 – β -коэффициенты, которые являются статистическими оценками параметров модели.

Оценка параметров множественной регрессии осуществляется по методу наименьших квадратов МНК). Значения β -коэффициентов находят как решение **системы нормальных уравнений**:

$$\begin{cases} r_{yx_1} = \beta_1 + r_{x_1x_2} \cdot \beta_2, \\ r_{yx_2} = r_{x_1x_2} \cdot \beta_1 + \beta_2. \end{cases} \quad (2.9)$$

Отсюда:

$$\beta_1 = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2} \quad \text{и} \quad \beta_2 = \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2}.$$

Преобразуя соотношение (2.8), получаем эконометрическую модель в натуральных переменных:

$$\widehat{y} = b_0 + b_1x_1 + b_2x_2, \quad (2.10)$$

где b_0 , b_1 и b_2 – статистические оценки параметров модели.

Статистические оценки параметров модели, построенной в натуральных переменных, определяются соотношениями:

$$b_j = \beta_j \cdot \frac{\sigma_y}{\sigma_{x_j}}, \quad \text{где } j = \overline{1,2}, \quad (2.11)$$

$$b_0 = \overline{y} - b_1\overline{x_1} - b_2\overline{x_2}. \quad (2.12)$$

Частичные коэффициенты эластичности эндогенного фактора для каждого экзогенного фактора X_j ($j = \overline{1,2}$) определяются по формуле:

$$\mathcal{E}_{x_j} = b_j \cdot \frac{\overline{x_j}}{\overline{y}}, \quad \text{где } j = \overline{1,2}. \quad (2.13)$$

Для определения влияния каждого экзогенного фактора на эндогенный фактор при условии, что другой экзогенный фактор остается постоянным, вычисляют **частные коэффициенты корреляции**:

$$r_{yx_1|x_2} = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{\sqrt{(1 - r_{yx_2}^2) \cdot (1 - r_{x_1x_2}^2)}}; \quad (2.14)$$

$$r_{yx_2|x_1} = \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1x_2}}{\sqrt{(1 - r_{yx_1}^2) \cdot (1 - r_{x_1x_2}^2)}}. \quad (2.15)$$

Взаимное влияние факторов X_1 и X_2 при условии, что эндогенный фактор Y не изменяется, определяет частный коэффициент корреляции:

$$r_{x_1x_2|y} = \frac{r_{x_1x_2} - r_{yx_1} \cdot r_{yx_2}}{\sqrt{(1 - r_{yx_1}^2) \cdot (1 - r_{yx_2}^2)}}. \quad (2.16)$$

Для проверки статистической значимости частных коэффициентов корреляции (2.14) – (2.16) применяют критерий Стьюдента. Для каждого из частных коэффициентов корреляции проверке подлежит основная гипотеза $H_0: r_{x_1x_2|y} = 0$ при альтернативной гипотезе $H_1: r_{x_1x_2|y} \neq 0$. Эмпирическое значение критерия Стьюдента находят из соотношения:

$$t_{yx_i|x_j} = \frac{|r_{yx_i|x_j}|}{S_{yx_i|x_j}}, \quad (2.17)$$

где $S_{yx_i|x_j}^2$ – дисперсия соответствующего частного коэффициента корреляции, которая вычисляется по формуле:

$$S_{yx_i|x_j}^2 = \frac{1 - r_{yx_i|x_j}^2}{n - m - 1}. \quad (2.18)$$

Критические точки распределения Стьюдента определяют по таблице для уровней значимости $\alpha = 0,05$ и $\alpha = 0,01$ при количестве степеней свободы $k = n - m - 1$. Если $t_{эмн.} < t_{0,05}$, то гипотезу H_0 нет оснований отвергнуть; если $t_{эмн.} > t_{0,01}$, то с надежностью 99 % корреляционная связь является существенной. Область $t_{0,05} < t_{эмн.} < t_{0,01}$ является областью неопределенности.

Коэффициент множественной корреляции определяется соотношением:

$$R_{yx_1x_2} = \sqrt{r_{yx_1} \beta_1 + r_{yx_2} \beta_2}. \quad (2.19)$$

Значимость уравнения множественной регрессии в целом проверяют с помощью F -критерия. Проверке подлежит основная гипотеза

$H_0: R^2 = 0$ при альтернативной гипотезе $H_1: R^2 \neq 0$. Эмпирическое значение критерия Фишера вычисляют по формуле:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m}, \quad (2.20)$$

где n – объем выборочной совокупности (количество наблюдений);
 m – количество объясняющих факторов, рассматриваемых в модели;
 R^2 – коэффициент детерминации.

Эмпирическое значение F -критерия сравнивают с табличными, которые определяют для уровней значимости $\alpha = 0,05$ и $\alpha = 0,01$ при количестве степеней свободы $k_1 = m$ и $k_2 = n - m - 1$. Если $F_{эмп.} < F_{0,05}$, то гипотезу H_0 нет оснований отклонить; если $F_{эмп.} > F_{0,01}$, то с надежностью 99 % совокупное влияние экзогенных факторов на эндогенный следует считать существенным. Область $F_{0,05} < F_{эмп.} < F_{0,01}$ является областью неопределенности.

Целесообразность присутствия в модели каждого из экзогенных факторов проверяют с помощью **частных F -критериев**:

$$F_{x_1} = \frac{R^2 - r_{yx_2}^2}{1 - R^2} \cdot \frac{n - m - 1}{1}; \quad (2.21)$$

$$F_{x_2} = \frac{R^2 - r_{yx_1}^2}{1 - R^2} \cdot \frac{n - m - 1}{1}. \quad (2.22)$$

Напомним, что одним из условий применения метода наименьших квадратов является неизменность дисперсии эндогенного фактора для разных частей выборочной совокупности. Это свойство называется **гомоскедастичностью**. Если различным частям выборочной совокупности присущи различные значения дисперсии результативного признака, то есть имеет место **гетероскедастичность**, то наблюдения являются неоднородными. В этом случае применение классического метода МНК дает оценки, которые являются неэффективными, смещенными и несостоятельными.

Для проверки выборочной совокупности на гомоскедастичность используется **тест Голдфелда – Квандта**. В качестве статистического критерия применяется критерий Фишера, по которому сравнивают дисперсии $D_1(Y)$ и $D_2(Y)$ для двух частей выборочной совокупности, противоположных по значениям кожного из экзогенных факторов (из конца и изначала вариационного ряда, упорядоченного по каждому из экзогенных факторов отдельно). Соответственно, проверке подлежит основная гипотеза $H_0: D_1(Y) - D_2(Y) = 0$ при альтернативной $H_1: D_1(Y) - D_2(Y) \neq 0$.

2.3. Пример выполнения типового задания № 2

Исходные данные. В табл. 2.2 приведены эмпирические данные о фондоемкости Y (грн/тыс. грн), затратах предприятия на устранение всех видов простоя X_1 (тыс. грн) и издержках на модернизацию и совершенствование техники и технологии X_2 (тыс. грн), полученные как результат статистический наблюдений для пятнадцати предприятий одной отрасли.

Таблица 2.2

Исходные данные задачи

№	Y	X_1	X_2
1	31,5	80,8	48,1
2	29,3	78,3	67,2
3	14,6	232,5	178,2
4	34,5	84,2	105,3
5	29,7	79,1	91,3
6	16,7	209,8	198,6
7	19,5	206,2	162,2
8	25,4	142,3	110,5
9	27,3	75,2	82,3
10	19,5	200,1	156,4
11	18,3	178,8	187,1
12	29,7	178,9	118,7
13	36,2	91,3	46,3
14	30,4	85,4	142,3
15	28,3	145,3	162,3

Рассматривая совокупность данных X_1 , X_2 и Y как трехмерную случайную величину, определим ее основные числовые характеристики. Необходимые вспомогательные вычисления представлены в табл. 2.3.

Таблица 2.3

Таблица вспомогательных вычислений

№	Y	X_1	X_2	Y^2	X_1^2	X_2^2	$Y \cdot X_1$	$Y \cdot X_2$	$X_1 \cdot X_2$
1	31,5	80,8	48,1	992,25	6 528,64	2 313,61	2 545,20	1 515,15	3 886,48
2	29,3	78,3	67,2	858,49	6 130,89	4 515,84	2 294,19	1 968,96	5 261,76
3	14,6	232,5	178,2	213,16	54 056,25	31 755,24	3 394,50	2 601,72	4 1431,5
4	34,5	84,2	105,3	1 190,25	7 089,64	11 088,09	2 904,90	3 632,85	8 866,26
5	29,7	79,1	91,3	882,09	6 256,81	8 335,69	2 349,27	2 711,61	7 221,83
6	16,7	209,8	198,6	278,89	44 016,04	39 441,96	3 503,66	3 316,62	41 666,28
7	19,5	206,2	162,2	380,25	42 518,44	26 308,84	4 020,90	3 162,90	33 445,64
8	25,4	142,3	110,5	645,16	20 249,29	12 210,25	3 614,42	2 806,70	15 724,15
9	27,3	75,2	82,3	745,29	5 655,04	6 773,29	2 052,96	2 246,79	6 188,96
10	19,5	200,1	156,4	380,25	40 040,01	24 460,96	3 901,95	3 049,80	31 295,64
11	18,3	178,8	187,1	334,89	31 969,44	35 006,41	3 272,04	3 423,93	33 453,48
12	29,7	178,9	118,7	882,09	32 005,21	14 089,69	5 313,33	3 525,39	21 235,43
13	36,2	91,3	46,3	1 310,44	8 335,69	2 143,69	3 305,06	1 676,06	4 227,19
14	30,4	85,4	142,3	924,16	7 293,16	20 249,29	2 596,16	4 325,92	12 152,42
15	28,3	145,3	162,3	800,89	21 112,09	26 341,29	4 111,99	4 593,09	23 582,19
Σ	390,9	2 068	1 856,8	10 819	333 257	265 034	49 181	44 557	289 639
Ср	26,06	137,88	123,78	721,24	22 217,11	17 668,94	3 278,70	2 970,50	19 309,28

Результаты вычисления основных числовых характеристик эмпирического распределения представлены в табл. 2.4.

Таблица 2.4

Основные числовые характеристики эмпирического распределения

Средняя выборки	Смещенная оценка среднего квадратического отклонения	Среднее выборочное отклонение (исправленное)	Парный коэффициент корреляции
$\bar{y} = 26,06$	$\sigma_y = 6,49$	$S_y = 6,72$	$r_{yx_1} = -0,8558$
$\bar{x}_1 = 137,88$	$\sigma_{x_1} = 56,62$	$S_{x_1} = 58,61$	$r_{yx_2} = -0,8125$
$\bar{x}_2 = 123,78$	$\sigma_{x_2} = 48,43$	$S_{x_2} = 50,13$	$r_{x_1x_2} = 0,8174$

Построим эконометрическую модель в стандартизированных переменных (2.8). Значения β -коэффициентов находим как решение системы уравнений (2.9), коэффициентами которых являются основные числовые характеристики выборочного распределения (см. табл. 2.4):

$$\begin{cases} -0,8558 = \beta_1 + 0,8174 \cdot \beta_2, \\ -0,8125 = 0,8174 \cdot \beta_1 + \beta_2, \end{cases} \Rightarrow \begin{cases} \beta_1 = -0,5776, \\ \beta_2 = -0,3404. \end{cases}$$

Итак, имеем эконометрическую модель в стандартизированных переменных:

$$\hat{t}_y = -0,5776 \cdot t_{x_1} - 0,3404 \cdot t_{x_2}.$$

Так как $|\beta_1| > |\beta_2|$, то можно утверждать, что расходы X_1 , связанные с устранением всех видов простоя, наиболее сильно влияют на фондоемкость Y производства.

Для построения уравнения регрессии в натуральных переменных по формулам (2.11) и (2.12) вычисляем МНК-оценки параметров модели:

$$b_1 = -0,5776 \cdot \frac{6,49}{56,62} = -0,0662;$$

$$b_2 = -0,3404 \cdot \frac{6,49}{48,43} = -0,0456;$$

$$b_0 = 26,06 + 0,0662 \cdot 137,88 + 0,0456 \cdot 123,78 = 40,8343.$$

В натуральных переменных эконометрическая модель имеет вид:

$$\hat{y} = 40,8343 - 0,0662 \cdot x_1 - 0,0456 \cdot x_2.$$

По формулам (2.13) вычисляем частичные коэффициенты эластичности:

$$\mathcal{E}_{x_1} = -0,0662 \cdot \frac{137,88}{26,06} \approx -0,35,$$

$$\mathcal{E}_{x_2} = -0,0456 \cdot \frac{123,78}{26,06} \approx -0,22.$$

Это означает, что при увеличении расходов на устранение всех видов простоя на 1 % среднее значение фондоемкости уменьшится на 0,35 %, а увеличение затрат на модернизацию на 1 % приведет к уменьшению среднего значения фондоемкости на 0,22 %.

По формулам (2.14) и (2.15) вычисляем частные коэффициенты корреляции:

$$r_{yx_1|x_2} = \frac{-0,8558 - (-0,8125) \cdot 0,8174}{\sqrt{(1 - (-0,8125)^2) \cdot (1 - 0,8174^2)}} \approx -0,5707;$$

$$r_{yx_2|x_1} = \frac{-0,8125 - (-0,8558) \cdot 0,8174}{\sqrt{(1 - (-0,8558)^2) \cdot (1 - 0,8174^2)}} \approx -0,3791.$$

По формуле (2.16) определим частный коэффициент корреляции, характеризующий взаимное влияние факторов X_1 и X_2 при постоянном значении признака Y :

$$r_{x_1x_2|y} = \frac{0,8174 - (-0,8558) \cdot (-0,8125)}{\sqrt{(1 - (-0,8558)^2) \cdot (1 - (-0,8125)^2)}} \approx 0,4048.$$

Связь между факторами является статистически значимой, но не очень тесной, поэтому частные коэффициенты корреляции мало отличаются от коэффициентов парной корреляции.

Проверим значимость частных коэффициентов корреляции по критерию Стьюдента. Для этого по формулам (2.18) определим дисперсию частных коэффициентов корреляции:

$$S_{r_{yx_1|x_2}}^2 = \frac{1 - (-0,5707)^2}{15 - 2 - 1} \approx 0,0561;$$

$$S_{r_{yx_2|x_1}}^2 = \frac{1 - (-0,3791)^2}{15 - 2 - 1} \approx 0,0714.$$

По формулам (2.17) вычислим эмпирические значения t -критерия для каждого из частных коэффициентов корреляции:

$$t_{yx_1|x_2} = \frac{|-0,5707|}{0,237} \approx 2,408; \quad t_{yx_2|x_1} = \frac{|-0,3791|}{0,267} \approx 1,419.$$

Для проверки основной гипотезы о статистической незначимости частных коэффициентов корреляции определяем критические точки распределения Стьюдента для уровней значимости $\alpha = 0,05$ и $\alpha = 0,01$ при количестве степеней свободы $k = 15 - 2 - 1 = 12$. Из справочной таблицы (см. табл. В.1 приложения В) находим: $t_{0,05} = 2,18$ и $t_{0,01} = 3,05$. Так как для фактора X_1 имеет место соотношение $t_{эмн.} > t_{0,05}$, то с надежностью 95 % корреляционная связь между затратами на устранение всех видов простой и фондоемкостью следует считать существенным. Напротив, для фактора X_2 имеем $t_{эмн.} < t_{0,05}$, то есть влияние расходов на модернизацию на фондоемкость можно считать статистически несущественным.

По формуле (2.19) вычисляем коэффициент множественной корреляции:

$$R_{yx_1x_2} = \sqrt{-0,8558 \cdot (-0,5776) + (-0,8125) \cdot (-0,3404)} \approx 0,8780.$$

Определим коэффициент детерминации: $R_{yx_1x_2}^2 = 0,8780^2 \approx 0,7709$.

Итак, изменчивость фондоемкости (Y) на 77,09 % определяется изменчивостью расходов на устранение всех видов простоя (X_1) и на модернизацию (X_2). Остальные 22,91 % обусловлены влиянием факторов, которые не учтены в данной модели.

Проверим значимость уравнения множественной регрессии в целом по критерию Фишера. По формуле (2.20) определяем эмпирическое значение критерия:

$$F = \frac{0,7709}{1 - 0,7709} \cdot \frac{15 - 2 - 1}{2} \approx 20,187.$$

Определим критические значения распределения Фишера для уровней значимости $\alpha = 0,05$ и $\alpha = 0,01$ при количестве степеней свободы $k_1 = 2$ и $k_2 = 12$. Из справочной таблицы находим $F_{0,05} = 3,88$ (см. табл. А.1 приложения А) и $F_{0,01} = 6,9$ (см. табл. Б. 1 приложения Б). Так как $F_{эмн.} > F_{0,01}$, то с надежностью 99 % основная гипотеза отклоняется в пользу альтернативной, то есть регрессия в целом является значимой.

Проверим, выполняется ли для выборочной совокупности условие гомоскедастичности по каждому из внешних факторов. Согласно условиям

теста Голдфельда – Квандта ранжируем остатки модели множественной регрессии отдельно по каждому из внешних факторов X_1 и X_2 . В табл. 2.5 приведены результаты вычислений, необходимых для проверки выборочной совокупности на наличие гетероскедастичности остатков по фактору X_1 . Упорядоченную выборку разделяем на три части и для ее первой и третьей частей (в порядке возрастания значения фактора X_1) построим уравнения регрессии. Получаем: $\hat{y} = -8,68 + 0,486 \cdot x_1$ и $\hat{y} = 56,259 - 0,182 \cdot x_1$. Для кожного из этих уравнений найдем сумму квадратов ошибок модели $\sum e_i^2$.

Таблица 2.5

Проверка отсутствия гетероскедастичности по фактору X_1

№	Y	X_1	\hat{y} по группам	e_i по группам	e_i^2 по группам
1	27,3	75,2	27,8737	-0,5737	0,3291
2	29,3	78,3	29,3806	-0,0806	0,0065
3	29,7	79,1	29,7695	-0,0695	0,0048
4	31,5	80,8	30,5958	0,9042	0,8175
5	34,5	84,2	32,2486	2,2514	5,0689
6	30,4	85,4	32,8319	-2,4319	5,9142
	36,2	91,3			$\sum e_i^2 = 12,141$
	25,4	142,3			
	28,3	145,3			
1	18,3	178,8	23,7608	-5,4608	29,8203
2	29,7	178,9	23,7426	5,9574	35,4904
3	19,5	200,1	19,8893	-0,3893	0,1516
4	19,5	206,2	18,7806	0,7194	0,5175
5	16,7	209,8	18,1263	-1,4263	2,0343
6	14,6	232,5	14,0004	0,5996	0,3596
					$\sum e_i^2 = 68,374$

Поскольку количество объектов в обеих подвыборках одинаково, то для вычисления эмпирического значения F -критерия определяем отношение большей из сумм квадратов ошибок к меньшей:

$$F = \frac{68,374}{12,141} = 5,63.$$

Находим критическое значение распределения Фишера для уровня значимости $\alpha = 0,05$ при количестве степеней свободы $k_1 = 4$ и $k_2 = 4$ (см. табл. А.1 приложения А): $F_{0,05}(4, 4) = 6,39$. Так как $F_{эмп.} < F_{0,05}$, то гипотезу H_0 об отсутствии гетероскедастичности по переменной X_1 нет оснований отклонить.

Теперь проверим гипотезу об отсутствии гетероскедастичности по переменной X_2 . Для первой и третьей подвыборок построим отдельные уравнения регрессии: $\hat{y} = 33,690 - 0,031 \cdot x_2$ и $\hat{y} = 45,742 - 0,151 \cdot x_2$. Для каждого из них найдем сумму квадратов ошибок модели $\sum e_i^2$ (табл. 2.6).

Таблица 2.6

Проверка отсутствия гетероскедастичности по фактору X_2

№	Y	X_2	\hat{y} по группам	e_i по группам	e_i^2 по группам
1	31,5	48,1	32,2005	-0,7005	0,4907
2	36,2	46,3	32,2562	3,9438	15,5536
3	29,3	67,2	31,6091	-2,3091	5,3321
4	27,3	82,3	31,1416	-3,8416	14,7582
5	29,7	91,3	30,8630	-1,1630	1,3526
6	34,5	105,3	30,4296	4,0704	16,5685
	25,4	110,5			$\sum e_i^2 = 54,056$
	29,7	118,7			
	30,4	142,3			
1	19,5	156,4	22,1574	-2,6574	7,0620
2	19,5	162,2	21,2828	-1,7828	3,1785
3	28,3	162,3	21,2677	7,0323	49,4526
4	14,6	178,2	18,8701	-4,2701	18,2337
5	18,3	187,1	17,5280	0,7720	0,5960
6	16,7	198,6	15,7939	0,9061	0,8211
					$\sum e_i^2 = 79,344$

По данным табл. 2.6 вычисляем эмпирическое значение статистики Фишера – Снедекора по переменной X_2 :

$$F = \frac{79,344}{54,056} = 1,47.$$

Критическое значение равно $F_{0,05}(4, 4) = 6,39$ (см. табл. А.1 приложения А). Сравнение этих значений показывает, что $F_{эмн.} < F_{0,05}$, следовательно, гетероскедастичности по переменной X_2 нет.

Проверим целесообразность присутствия каждого из внешних факторов в эконометрической модели с помощью частных F -критериев. Вычислим их эмпирические значения по формулам (2.21) и (2.22):

$$F_{x_1} = \frac{0,7709 - (-0,8125)^2}{1 - 0,7709} \cdot \frac{15 - 2 - 1}{1} \approx 5,801,$$

$$F_{x_2} = \frac{0,7709 - (-0,8558)^2}{1 - 0,7709} \cdot \frac{15 - 2 - 1}{1} \approx 2,017.$$

Сравним эти эмпирические значения с критическими точками распределения Фишера – Снедекора для уровней значимости $\alpha = 0,05$ и $\alpha = 0,01$ при числе степеней свободы $k_1 = 1$ и $k_2 = 12$. Так, $F_{0,05} = 4,75$ (см. табл. А.1 приложения А) и $F_{0,01} = 9,3$ (см. табл. Б.1 приложения Б). Так как для фактора X_1 имеем: $F_{эмн.} > F_{0,05}$, то с надежностью 95 % влияние фактора X_1 на признак Y нельзя считать несущественным. Для фактора X_2 : $F_{эмн.} < F_{0,05}$, то есть гипотезу об отсутствии влияния этого фактора на признак Y нет оснований отклонить и его наличие в модели является нецелесообразной. Это совпадает с выводами относительно значимости корреляционной связи по критерию Стьюдента.

Вывод. При анализе фондоемкости производства целесообразно использовать модель парной регрессии, которая учитывает только влияние такого фактора, как затраты на устранение всех видов простоя. В натуральных переменных эта модель имеет вид:

$$\hat{y} = 39,585 - 0,0981 \cdot x_1.$$

МНК-оценки по этой модели являются несмещенными, состоятельными и эффективными.

Задание 3. Моделирование временного ряда: анализ и прогноз

3.1. Задания для самостоятельного выполнения

По эмпирическим данным (табл. 3.1) об изменениях объемов продажи промышленной продукции:

построить уравнение линии тренда в предположении линейной корреляционной зависимости и определить сезонную составляющую временного ряда;

построить многофакторную модель временного ряда с использованием dummy-переменных, которая одновременно учитывает и трендовую, и сезонную составляющие уровней эндогенного фактора;

сделать прогноз по обеим моделям и сравнить результаты.

Таблица 3.1

Исходные данные для построения моделей динамики

№ варианта	Факторы	Значения факторов							
		3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10
1	t	1	2	3	4	5	6	7	8
	y	7,1	6,5	6,3	6,5	7,5	7,4	7,5	9,1
	t	9	10	11	12	13	14	15	16
	y	8,5	8,3	8,5	10,5	9,4	9,5	10,3	11,5
2	t	1	2	3	4	5	6	7	8
	y	4,1	3,5	3,3	3,5	4,8	4,4	4,5	5,1
	t	9	10	11	12	13	14	15	16
	y	6,5	5,3	5,5	6,5	7,8	6,5	6,8	7,5
3	t	1	2	3	4	5	6	7	8
	y	6,2	5,3	4,7	5,3	6,7	6,6	6,8	7,7
	t	9	10	11	12	13	14	15	16
	y	9,8	8,0	8,3	9,8	11,7	9,8	10,2	11,3
4	t	1	2	3	4	5	6	7	8
	y	3,2	4,3	3,7	3,3	3,7	4,7	4,6	4,7
	t	9	10	11	12	13	14	15	16
	y	5,4	6,8	5,6	5,8	6,8	8,2	6,8	7,1

Продолжение табл. 3.1

1	2	3	4	5	6	7	8	9	10
5	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	6,4	8,6	7,4	6,6	7,4	9,4	9,2	9,5
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	10,7	13,7	11,1	11,6	13,7	16,4	13,7	14,3
6	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	5,9	5,2	5,9	8,2	7,4	7,6	8,6	10,9
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	9,2	10,9	13,1	10,9	11,4	12,6	15,6	12,1
7	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	4,6	6,2	5,3	4,7	5,3	7,3	6,7	6,8
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	7,7	9,8	8,0	8,3	9,8	11,8	9,8	10,3
8	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	8,2	11,4	10,1	9,0	9,5	14,2	12,0	12,2
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	13,9	17,7	14,4	15,0	17,7	21,2	18,3	18,0
9	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	7,8	6,9	7,4	10,9	9,2	9,5	10,7	13,7
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	11,1	11,6	13,7	16,4	14,1	13,9	14,7	19,5
10	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	5,3	7,3	6,7	6,8	7,7	9,8	8,0	8,3
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	9,8	11,8	9,8	10,3	11,3	14,0	10,9	11,3
11	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	9,5	14,2	12,0	12,2	13,9	17,7	14,4	15,0
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	17,7	21,2	18,3	18,0	19,1	25,2	19,6	20,3
12	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	4,9	6,9	6,0	5,4	5,7	8,5	7,2	7,3
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	8,7	10,6	9,0	9,1	10,3	12,7	11,0	10,8

Продолжение табл. 3.1

1	2	3	4	5	6	7	8	9	10
13	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	5,3	7,8	6,6	6,8	8,0	9,8	8,3	8,4
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	9,5	11,7	10,1	9,9	10,5	13,5	10,8	11,2
14	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	3,7	3,5	3,6	5,4	4,6	4,7	5,3	6,8
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	5,8	5,9	6,6	8,2	7,1	6,9	7,4	9,5
15	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	5,7	8,6	7,4	7,6	8,4	10,9	9,2	9,4
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	10,6	13,1	11,3	11,1	11,8	15,1	12,1	12,5
16	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	3,8	4,2	6,1	5,3	5,0	5,1	7,7	6,7
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	6,8	7,6	9,8	8,3	8,5	9,5	11,8	10,2
17	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	5,7	8,5	7,2	7,3	8,7	10,6	9,0	9,1
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	10,3	12,7	11,0	10,8	11,4	15,1	11,8	12,2
18	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	3,8	4,2	6,1	5,3	5,0	5,1	7,7	6,6
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	6,8	7,5	9,8	8,3	8,4	9,5	11,7	10,1
19	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	8,7	9,6	13,9	12,0	11,3	11,7	17,5	15,1
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	15,4	17,1	22,3	18,9	19,2	21,6	26,7	23,0
20	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	5,1	7,7	6,5	6,5	7,2	9,3	8,3	8,4
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	9,5	11,7	10,1	10,2	10,9	13,5	10,8	11,2

1	2	3	4	5	6	7	8	9	10
21	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	5,9	5,2	5,7	8,6	7,2	7,2	8,1	10,4
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	9,2	9,4	10,6	13,1	11,3	11,4	12,2	15,1
22	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	7,8	7,6	11,0	9,5	8,4	9,3	13,9	11,7
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	11,7	13,1	16,9	15,0	15,2	17,1	21,2	18,3
23	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	11,8	11,5	16,7	14,4	13,6	14,0	21,0	18,1
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	18,5	20,6	26,7	22,6	23,0	25,9	32,1	27,6
24	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	4,3	4,7	6,8	5,9	5,5	5,7	8,6	7,4
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	7,6	8,4	10,9	9,2	9,4	10,6	13,1	11,3
25	<i>t</i>	1	2	3	4	5	6	7	8
	<i>y</i>	11,7	17,5	15,1	15,4	17,1	22,3	18,9	19,2
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>y</i>	21,6	26,7	23,0	22,6	24,0	30,9	24,7	25,6

3.2. Содержательная постановка задачи и ее математическая модель

Временным рядом, или **рядом динамики** называется совокупность значений какого-либо показателя, описывающая определенный экономический или социальный процесс за несколько последовательных промежутков или моментов времени. Основными задачами исследования временных рядов является определение количественных соотношений, отражающих вклад всей совокупности экзогенных факторов, которые влияют на исследуемый процесс и изменяются во времени, в значение эндогенного фактора, а также дальнейшее использование полученной модели для прогнозирования развития исследуемых процессов.

Прогноз – это вероятностное, научно обоснованное суждение относительно перспектив и возможных состояний того или иного объекта в будущем и/или путей и сроков достижения будущих состояний. Прогнозирование предусматривает следующие этапы: сбор первичной информации и ее обработку с целью построения математической модели исследуемого процесса. В зависимости от момента в будущем, для которого строится прогноз, выделяют краткосрочные, среднесрочные и долгосрочные прогнозы. В моделях временных рядов, которые в дальнейшем будут использоваться для построения краткосрочных прогнозов, под циклическим компонентом понимают, например, сезонные колебания. В моделях, предназначенных для построения долгосрочных прогнозов, кроме сезонных колебаний необходимо также учитывать циклические процессы, имеющие длительный период. Следует отметить, что наибольший период, на который можно делать прогноз на базе построенной модели, соотносится с периодом наблюдений, результаты которых применялись во время ее построения, как 1:3.

Хотя временной ряд можно рассматривать как одну из разновидностей парной стохастической зависимости, где роль экзогенного фактора играет время, построение математической модели временных рядов имеет определенные особенности. Эти особенности обусловлены как структурой временного ряда, так и тем, что для временных рядов часто нарушается одно из условий теоремы Гаусса – Маркова, на которой базируется метод наименьших квадратов, а именно – наблюдается автокорреляция остатков модели.

Рассмотрим структуру временного ряда. Значение эндогенного фактора, которое соответствует определенному моменту времени, называется **уровнем ряда**. Для рядов динамики значение эндогенного фактора задают через одинаковые промежутки времени, поэтому последовательность значений может отображаться их порядковым номером. Существуют **интервальные** и **моментные временные ряды**. Для интервального ряда предоставляются суммарные показатели за определенный период времени (например, за сутки, месяц, год). Особенностью этого ряда является то, что его уровни можно складывать. Для моментных рядов значения уровня рассматриваются в определенный момент времени. Для таких рядов сложение уровней не имеет смысла, однако их разница дает информацию об изменении показателей на протяжении определенного периода времени.

В общем случае временной ряд содержит как **детерминированную** (обусловленную действием экзогенных факторов, влияние которых исследуется в данной модели), так и **случайную составляющие (ошибки модели)**. Если можно предположить, что детерминированная и случайная составляющие аддитивные, то для определения уровня временного ряда применимо соотношение:

$$y_t = f(t) + \varepsilon_t, \quad t = \overline{1, n}, \quad (3.1)$$

где y_t – уровень временного ряда, то есть значение эндогенного фактора в момент времени t (или за период), которое является результатом влияния всей совокупности экзогенных факторов;

$f(t)$ – детерминированная составляющая;

ε_t – случайная составляющая, которая, согласно требованиям теоремы Гаусса – Маркова, характеризуется математическим ожиданием $M(\varepsilon_t) = 0$ и дисперсией $D(\varepsilon_t) = \sigma^2$;

n – длина ряда (количество уровней).

В свою очередь, можно предполагать, что детерминированная составляющая временного ряда формируется под влиянием нескольких экзогенных факторов, которые можно разделить на две группы: факторы, определяющие тенденцию, то есть **тренд**, и те факторы, которые определяют **циклические колебания**. В большинстве исследований это сезонные колебания. Модель временного ряда, в которой уровень эндогенного фактора определяется как сумма приведенных выше компонентов, называется **аддитивной**. В матричной форме она имеет вид:

$$\mathbf{Y} = \mathbf{T} + \mathbf{S} + \mathbf{E}, \quad (3.2)$$

где \mathbf{Y} – матрица эмпирических значений эндогенного фактора;

\mathbf{T} – матрица расчетных значений трендовой компоненты;

\mathbf{S} – матрица расчетных значений сезонной компоненты;

\mathbf{E} – матрица ошибок, определяет расхождение между эмпирическими значениями экзогенного фактора и его значениями, рассчитанными в соответствии с математической моделью.

Исследование структуры временного ряда предполагает его декомпозицию. Одним из возможных путей определения трендовой составляющей

является построение уравнения парной регрессии по эмпирическим значениями эндогенного фактора. Тогда сезонная составляющая для определенного периода (**сезона**) определяется как среднее значение разностей между эмпирическими значениями эндогенного фактора и его значениями, рассчитанными по уравнению парной регрессии.

Построение многофакторной эконометрической модели временного ряда с применением дополнительных переменных позволяет разделять трендовую и циклическую составляющие в пределах одной модели. В этом случае в состав экзогенных факторов кроме переменной, которая определяет порядковый номер уровня, включают ***dummy-переменные***. Эти переменные определяют принадлежность уровня эндогенного фактора определенному периоду, то есть отражают влияние факторов, которые являются качественными по своей природе. Например, с помощью *dummy-переменных* определяют, принадлежит ли данное измерение определенному кварталу, если в модели временного ряда учитываются сезонные колебания. В общем случае *dummy-переменным* приписывают значения 0 (нет) или 1 (да) в зависимости от того, выполняется ли конкретное условие.

Количество фиктивных переменных зависит от количества уровней, принадлежащих одному периоду. Так, для полной характеристики динамики любого процесса достаточно выбрать *dummy-переменные*, количество которых на единицу меньше количества уровней, которые целесообразно считать принадлежащими одному периоду.

3.3. Пример выполнения типового задания № 3

Исходные данные. Временной ряд содержит сведения о ежеквартальной прибыли фирмы (тыс. грн) в течение четырех лет (табл. 3.2).

Таблица 3.2

Квартальная прибыль фирмы

Годы	Кварталы			
	I	II	III	IV
2009	38,2	52,0	61,0	50,5
2010	48,5	61,8	78,6	67,2
2011	58,2	70,4	91,2	78,4
2012	64,0	72,6	98,8	90,6

Для того чтобы сделать предположение относительно структуры временного ряда, по исходным данным (см. табл. 3.2) построим график изменения прибыли фирмы во времени (рис. 3.1). Следует отметить, что исходные данные представляют собой шестнадцать уровней значений эндогенного фактора.

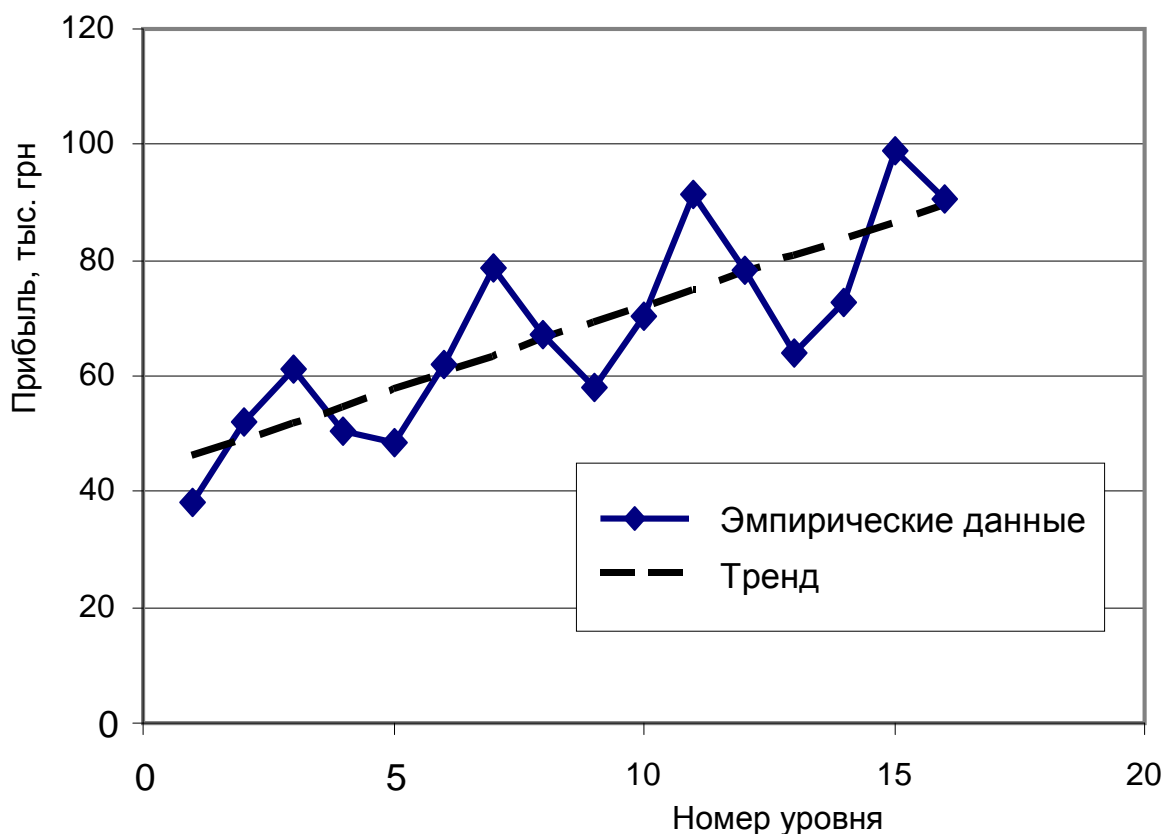


Рис. 3.1. Динамика роста прибыли фирмы по кварталам

На рис. 3.1 видно, что в течение исследуемого периода существует тенденция роста прибыли фирмы, для аппроксимации которой можно выбрать линейную функцию. Кроме того, происходят сезонные колебания уровня прибыли относительно линии тренда, периодом которых следует считать год. Согласно такому представлению математическая модель процесса имеет вид:

$$\hat{f}(t) = \underbrace{b_0 + b_1 \cdot t}_{\text{тренд}} + \underbrace{\hat{s}_i}_{\text{сезонная составляющая}}, \quad i = \overline{1,4}, \quad (3.3)$$

где $\hat{f}(t)$ – значение детерминированной составляющей уровня эндогенного фактора, вычисленное в соответствии с принятой моделью;

b_0, b_1 – параметры модели линейной регрессии;

\hat{s}_i – значение сезонной составляющей в соответствии с моделью.

Напомним, что по методу наименьших квадратов статистическая оценка коэффициента регрессии в случае линейной аппроксимации осуществляется по формуле:

$$b_1 = \frac{\overline{y \cdot t} - \bar{y} \cdot \bar{t}}{\overline{t^2} - (\bar{t})^2}, \quad (3.4)$$

где \bar{y} – выборочная средняя эндогенного фактора;

\bar{t} – выборочная средняя экзогенного фактора (номера уровня);

$\overline{y \cdot t}$ – выборочная средняя произведения факторов;

$\overline{t^2}$ – выборочная средняя квадрата экзогенного фактора.

Соответственно, статистическая оценка свободного члена уравнения регрессии в этом случае равна:

$$b_0 = \bar{y} - b_1 \cdot \bar{t}. \quad (3.5)$$

Методика построения уравнения парной регрессии подробно описана в задании 1. Приведем результаты промежуточных вычислений, необходимых для оценки параметров уравнения линии тренда (табл. 3.3).

Таблица 3.3

Результаты промежуточных вычислений для оценки параметров линейной модели парной регрессии

Год	Квартал	t	y	$y \cdot t$	t^2	$b_0 + b_1 t$	$y - (b_0 + b_1 t)$
1	2	3	4	5	6	7	8
2012	I	1	38,2	38,2	1	45,94	-7,74
	II	2	52	104,0	4	48,83	3,17
	III	3	61	183,0	9	51,72	9,28
	IV	4	50,5	202,0	16	54,61	-4,1
2013	I	5	48,5	242,5	25	57,51	-9,01
	II	6	61,8	370,8	36	60,40	1,4
	III	7	78,6	550,2	49	63,29	-5,31
	IV	8	67,2	537,6	64	66,18	1,02

1	2	3	4	5	6	7	8
2014	I	9	58,2	523,8	81	69,07	-10,87
	II	10	70,4	704,0	100	71,963	-1,56
	III	11	91,2	1003,2	121	74,85	16,35
	IV	12	78,4	940,8	144	77,75	0,65
2015	I	13	64	832,0	169	80,64	-16,64
	II	14	72,6	1016,4	196	83,53	-10,93
	III	15	98,8	1482,0	225	86,42	12,38
	IV	16	90,6	1449,6	256	89,31	1,29
Выборочная средняя		8,5	67,62	636,26	93,5		

Подставив значения соответствующих выборочных средних (их значения приведены в последней строке столбиков 3 – 6 табл. 3.2), по формуле (3.4) вычисляем коэффициент регрессии:

$$b_1 = \frac{636,26 - 67,62 \cdot 8,5}{93,5 - (8,5)^2} = 2,89.$$

Подставляем найденное значение в формулу (3.5) и вычисляем свободный член уравнения регрессии:

$$b_0 = 67,62 - 2,89 \cdot 8,5 = 43,05.$$

Для определения сезонной составляющей эконометрической модели временного ряда сначала по уравнению регрессии для каждого номера уровня вычисляем значение трендовой составляющей эндогенного фактора. Они записаны в столбце 7 табл. 3.2. Затем определяем разность между эмпирическим значением эндогенного фактора и его значением, рассчитанным по уравнению тренда. Эти результаты записаны в столбце 8 табл. 3.2. Сезонные составляющие для каждого квартала определяют как среднее арифметическое разностей, соответствующих данному кварталу.

Так, для I квартала имеем:

$$s_I = \frac{-7,74 - 9,01 - 10,87 - 16,64}{4} = -11,06;$$

для II квартала:

$$s_{II} = \frac{3,17 + 1,4 - 1,56 - 10,93}{4} = -1,98;$$

для III квартала:

$$s_{III} = \frac{9,28 - 5,31 + 16,35 + 12,38}{4} = 13,33;$$

для IV квартала:

$$s_{IV} = \frac{-4,1 + 1,02 + 0,65 + 1,29}{4} = -0,29.$$

Окончательно математическую модель, описывающую изменение прибыли фирмы во времени, можно записать в виде:

$$\hat{y} = 43,05 + 2,89 \cdot t + s_i, \quad i = \overline{1,4}. \quad (3.6)$$

Используем полученную модель для прогноза. Поскольку общая продолжительность исследуемого периода составляет 16 периодов, то максимальная отдаленность прогноза равна 5-ти периодам. Следовательно, для определения прибыли фирмы в I квартале 2017 подставим в уравнение (3.6) значения $t = 21$, а $i = 1$. Прибыль может составить:

$$\hat{y} = 43,05 + 2,89 \cdot 21 - 11,06 = 92,67 \text{ (тыс. грн).}$$

Теперь построим математическую модель задачи как модель многофакторной регрессии, экзогенными факторами которой является номер уровня t и бинарные дитту-переменные d_1 , d_2 и d_3 , с помощью которых будем учитывать номер квартала. Так, первому кварталу любого года соответствует совокупность значений $d_1 = d_2 = d_3 = 0$, второму – совокупность значений $d_1 = 1, d_2 = d_3 = 0$, третьему – значения $d_2 = 1, d_1 = d_3 = 0$, а четвертому – значения $d_3 = 1, d_1 = d_2 = 0$. Математическую модель исследуемого процесса можно представить в таком виде:

$$\hat{y}(t) = b_0 + b_1 \cdot t + b_2 \cdot d_1 + b_3 d_2 + b_4 d_3, \quad (3.7)$$

где \hat{y} – значение детерминированной составляющей уровня эндогенного фактора, вычислено в соответствии с принятой моделью;

b_1 – коэффициент регрессии, учитывающий влияние номера уровня;
 b_2, b_3, b_4 – коэффициенты регрессии, учитывающие влияние номера квартала;
 b_0 – свободный член уравнения регрессии.

Принцип построения эконометрической модели, которая содержит dummy-переменные, не отличается от того, который применялся в задаче 2 для построения многофакторной модели с натуральными переменными. Статистические оценки параметров модели (3.7) по методу МНК получаем, решая систему нормальных уравнений:

$$\begin{cases} b_0 + b_1 \cdot \bar{t} + b_2 \cdot \bar{d}_1 + b_3 \cdot \bar{d}_2 + b_4 \cdot \bar{d}_3 = \bar{y}; \\ b_0 \cdot \bar{t} + b_1 \cdot \bar{t}^2 + b_2 \cdot \overline{d_1 \cdot t} + b_3 \cdot \overline{d_2 \cdot t} + b_4 \cdot \overline{d_3 \cdot t} = \overline{y \cdot t}; \\ b_0 \cdot \bar{d}_1 + b_1 \cdot \overline{t \cdot d_1} + b_2 \cdot \overline{d_1^2} + b_3 \cdot \overline{d_2 \cdot d_1} + b_4 \cdot \overline{d_3 \cdot d_1} = \overline{y \cdot d_1}; \\ b_0 \cdot \bar{d}_2 + b_1 \cdot \overline{t \cdot d_2} + b_2 \cdot \overline{d_1 \cdot d_2} + b_3 \cdot \overline{d_2^2} + b_4 \cdot \overline{d_3 \cdot d_2} = \overline{y \cdot d_2}; \\ b_0 \cdot \bar{d}_3 + b_1 \cdot \overline{t \cdot d_3} + b_2 \cdot \overline{d_1 \cdot d_3} + b_3 \cdot \overline{d_2 \cdot d_3} + b_4 \cdot \overline{d_3^2} = \overline{y \cdot d_3}. \end{cases} \quad (3.8)$$

Поскольку мультиколлинеарность между экзогенными факторами отсутствует, то есть определитель основной матрицы системы (определитель, составленный из коэффициентов при неизвестных) отличен от нуля, то для ее решения можно применить формулы Крамера:

$$b_j = \frac{\Delta b_j}{\Delta}; \quad j = \overline{0, m}, \quad (3.9)$$

где j – индекс параметров многофакторной модели, для которого вычисляется коэффициент регрессии;

m – общее количество экзогенных факторов, содержится в многофакторной модели;

Δ – определитель основной матрицы системы;

Δ_j – определитель, получаемый из определителя основной матрицы системы путем замены его j -го столбца, состоящего из коэффициентов при j -м неизвестном, на столбец свободных членов системы уравнений.

Частично необходимые вычисления уже были проведены в процессе построения линии тренда парной модели линейной регрессии, а именно – определены выборочные средние \bar{y} , \bar{t} , $\overline{y \cdot t}$ и $\overline{t^2}$ (см. табл. 3.3). Значения факторов, по которым осуществляются вычисления МНК-оценок параметров системы линейных алгебраических уравнений (3.8), и результаты некоторых вспомогательных вычислений представлены в виде таблицы (табл. 3.4).

Последняя строка табл. 3.4 (значения в этой строке выделены жирным шрифтом) содержит значения средних, рассчитанные для тех величин, которые вынесены в заголовки соответствующих столбцов. Значения средних получены как результат деления суммы по столбику (эти суммы записаны в предпоследней строке) на количество измерений.

Таблица 3.4

Некоторые результаты промежуточных вычислений для оценки параметров линейной модели множественной регрессии

№	y	t	d_1	d_2	d_3	$y d_1$	$y d_2$	$y d_3$
1	38,2	1	0	0	0	0	0	0
2	52	2	1	0	0	52	0	0
3	61	3	0	1	0	0	61	0
4	50,5	4	0	0	1	0	0	50,5
5	48,5	5	0	0	0	0	0	0
6	61,8	6	1	0	0	61,8	0	0
7	78,6	7	0	1	0	0	78,6	0
8	67,2	8	0	0	1	0	0	67,2
9	58,2	9	0	0	0	0	0	0
10	70,4	10	1	0	0	70,4	0	0
11	91,2	11	0	1	0	0	91,2	0
12	78,4	12	0	0	1	0	0	78,4
13	64	13	0	0	0	0	0	0
14	72,6	14	1	0	0	72,6	0	0
15	98,8	15	0	1	0	0	98,8	0
16	90,6	16	0	0	1	0	0	90,6
Σ	1082	136	4	4	4	256,8	329,6	286,7
Ср.	67,62	8,5	0,25	0,25	0,25	16,05	20,6	17,92

В дополнение к уже известным данным о средних выборки получаем, что $\overline{d_1^2} = \overline{d_2^2} = \overline{d_3^2} = 0,25$; $\overline{t \cdot d_1} = 2$; $\overline{t \cdot d_2} = 2,25$; $\overline{t \cdot d_3} = 2,5$; $\overline{d_1 \cdot d_2} = \overline{d_1 \cdot d_3} = \overline{d_2 \cdot d_3} = 0$.

Подставим рассчитанные значения средних в систему нормальных уравнений (3.8) и получим:

$$\begin{cases} b_0 + b_1 \cdot 8,5 + b_2 \cdot 0,25 + b_3 \cdot 0,25 + b_4 \cdot 0,25 = 67,62; \\ b_0 \cdot 8,5 + b_1 \cdot 93,5 + b_2 \cdot 2 + b_3 \cdot 2,25 + b_4 \cdot 2,5 = 636,26; \\ b_0 \cdot 0,25 + b_1 \cdot 2 + b_2 \cdot 0,25 + b_3 \cdot 0 + b_4 \cdot 0 = 16,05; \\ b_0 \cdot 0,25 + b_1 \cdot 2,25 + b_2 \cdot 0 + b_3 \cdot 0,25 + b_4 \cdot 0 = 20,6; \\ b_0 \cdot 0,25 + b_1 \cdot 2,5 + b_2 \cdot 0 + b_3 \cdot 0 + b_4 \cdot 0,25 = 17,92. \end{cases} \quad (3.10)$$

По коэффициентам при неизвестных в системе (3.10) запишем и вычислим определитель основной матрицы системы. Он равен:

$$\Delta = \begin{vmatrix} 1 & 8,5 & 0,25 & 0,25 & 0,25 \\ 8,5 & 93,5 & 2 & 2,25 & 2,5 \\ 0,25 & 2 & 0,25 & 0 & 0 \\ 0,25 & 2,25 & 0 & 0,25 & 0 \\ 0,25 & 2,5 & 0 & 0 & 0,25 \end{vmatrix} = 0,07812.$$

Заменяв первый столбец определителя основной матрицы системы на столбец свободных членов, получим определитель Δ_0 :

$$\Delta_0 = \begin{vmatrix} 67,62 & 8,5 & 0,25 & 0,25 & 0,25 \\ 636,26 & 93,5 & 2 & 2,25 & 2,5 \\ 16,05 & 2 & 0,25 & 0 & 0 \\ 20,6 & 2,25 & 0 & 0,25 & 0 \\ 17,92 & 2,5 & 0 & 0 & 0,25 \end{vmatrix} = 2,6616.$$

В соответствии с формулами (3.9) имеем:

$$b_0 = \frac{\Delta b_0}{\Delta} = \frac{2,6616}{0,07812} = 34,07.$$

Аналогично вычисляем другие определители:

$$\Delta_1 = 0,2026; \quad \Delta_2 = 0,7329; \quad \Delta_3 = 1,9521; \quad \Delta_4 = 0,9116$$

и по формулам (3.9) вычисляем другие параметры модели:

$$b_1 = 2,5939; \quad b_2 = 9,3818; \quad b_3 = 24,9891; \quad b_4 = 11,6695.$$

Таким образом, модель временного ряда как модель многофакторной регрессии, которая одновременно учитывает и трендовую, и сезонную составляющие, можно записать в виде:

$$\hat{y}(t) = 34,07 + 2,59 \cdot t + 9,38 \cdot d_1 + 24,99d_2 + 11,67 \cdot d_3. \quad (3.11)$$

Рассчитаем значения уровней эндогенного фактора по модели многофакторной регрессии (3.11) и сравним их с эмпирическими данными (рис. 3.2). Видно, что вычисления по этой модели дают результаты, которые хорошо совпадают с эмпирическими данными.

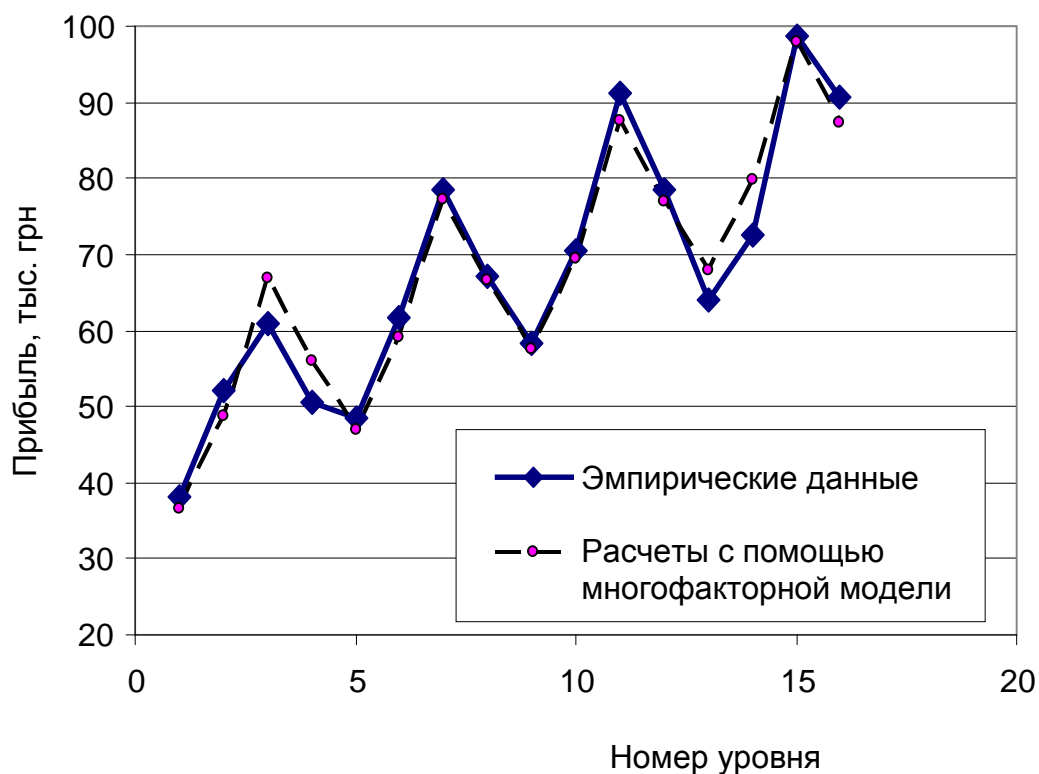


Рис. 3.2. Динамика роста прибыли фирмы по кварталам

Сравним качество построенных моделей. Для этого определим средние ошибки аппроксимации по каждой из моделей. Методика расчета рассмотрена в задании 1.

По результатам, приведенным в табл. 3.5, видно, что средняя ошибка аппроксимации при расчетах с использованием многофакторной модели меньше, чем по модели декомпозиции, но эти расхождения не являются существенными.

Таблица 3.5

Расчет ошибок аппроксимации по моделям динамики

Исходные данные		Модель декомпозиции		Многофакторная модель	
t	y_t	\hat{y}_t	$\left \frac{y_t - \tilde{y}_t}{y_t} \right \cdot 100\%$	\hat{y}_t	$\left \frac{y_t - \tilde{y}_t}{y_t} \right \cdot 100\%$
1	38,2	34,88	8,69	36,66	4,03
2	52	46,85	9,90	48,63	6,48
3	61	65,05	6,64	66,83	9,56
4	50,5	54,32	7,56	56,10	11,09
5	48,5	46,44	4,25	47,02	3,05
6	61,8	58,41	5,49	58,99	4,55
7	78,6	76,61	2,53	77,19	1,79
8	67,2	65,88	1,96	66,46	1,10
9	58,2	58,00	0,34	57,38	1,41
10	70,4	69,97	0,61	69,35	1,49
11	91,2	88,17	3,32	87,55	4,00
12	78,4	77,44	1,22	76,82	2,02
13	64	69,56	8,69	67,74	5,84
14	72,6	81,53	12,30	79,71	9,79
15	98,8	99,73	0,94	97,91	0,90
16	90,6	89,00	1,77	87,18	3,77
Сума			76,22		70,88
Средняя			4,76		4,43

Сравним результаты прогнозирования по обеим моделям. Для этого с помощью модели многофакторной регрессии сделаем прогноз

относительно прибыли компании в I квартале 2017, то есть для того же уровня, для которого ранее осуществлялся прогноз с помощью модели парной регрессии.

Подставим в уравнение (3.11) значения эндогенных переменных, для которых рассчитывается прогноз: номер уровня $t = 21$ и думми-переменные, которые отвечают первому кварталу $d_1 = d_2 = d_3 = 0$. Получаем значения прогноза на основе эконометрической модели многофакторной регрессии:

$$\hat{y}(t) = 34,07 + 2,59 \cdot 21 + 9,38 \cdot 0 + 24,99 \cdot 0 + 11,67 \cdot 0 = 88,54 \text{ (тыс. грн).}$$

Сравним полученные данные с результатом расчета по модели парной регрессии. Абсолютное расхождение между прогнозами, которые рассчитаны по моделям парной регрессии и многофакторной регрессии, составляет:

$$\Delta \hat{y} = 92,67 - 88,54 = 4,13 \text{ (тыс. грн).}$$

Это соответствует относительному расхождению между двумя прогнозами:

$$\frac{\Delta \hat{y}}{\hat{y}} \cdot 100 = \frac{4,13}{88,54} \cdot 100 = 4,67 \text{ (\%).}$$

Вывод. Исследования рядов динамики как с помощью эконометрической модели многофакторной регрессии, так и с помощью модели парной регрессии, которая предусматривает декомпозицию на тренд и сезонную составляющую, показали, что обе модели дают близкие результаты, поэтому обе они могут использоваться для анализа особенностей исследуемого процесса, а также для построения прогнозов на их основе. Однако период, на который можно распространять прогноз как по одной, так и по другой модели, не должен превышать пяти кварталов. Это обусловлено тем, что количество уровней эндогенного фактора, которое использовалось при построении эконометрических моделей, составляло только 16 значений.

Задание 4. Исследование автокорреляции остатков временного ряда

4.1. Задания для самостоятельного выполнения

По эмпирическим данным, которые характеризуют изменения объемов продаж промышленной продукции во времени (табл. 4.1):

построить с помощью метода МНК уравнение тренда в предположении линейной модели и проверить наличие автокорреляции остатков модели по критерию Дарбина – Уотсона;

определить оценки параметров модели с помощью обобщенного метода наименьших квадратов;

провести сравнительный анализ обеих моделей и сделать выводы.

Таблица 4.1

Исходные данные для исследования автокорреляции

№ варианта	Факторы	Значения факторов							
		3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10
1	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	8,1	6,5	6,3	6,5	7,5	7,4	7,5	10,1
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	8,5	8,3	8,5	9,5	9,4	9,5	10,5	10,3
2	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	10,1	8,8	8,8	11,6	10,4	10,1	9	12,1
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	10,8	10,8	13,6	12,4	12,1	11	12,8	12,8
3	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	5,6	6,2	6,5	6,6	5,5	6	6,5	7,6
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	8,2	8,5	8,6	7,5	8	8,5	10,2	10,5
4	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	8,8	7,8	9,3	9,4	9,9	10,1	8,8	10,8
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	9,8	11,3	11,4	11,9	12,1	10,8	11,8	13,3
5	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	26	28	25	28	30	32	35	28
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	30	27	30	32	34	37	32	29

Продолжение табл. 4.1

1	2	3	4	5	6	7	8	9	10
6	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	7,4	7,5	9,8	9	8,5	9,6	7,6	9,4
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	9,5	11,8	11	10,5	11,6	9,6	11,5	13,8
7	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	8,8	8,5	9,2	9,6	9,8	9,9	9,5	10,8
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	10,5	11,2	11,6	11,8	11,9	11,5	12,5	13,2
8	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	6,6	4,3	5,2	7,4	5,8	6,4	5,7	8,6
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	6,3	7,2	9,4	7,8	8,4	7,7	8,3	9,2
9	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	7,5	6,5	6,8	6,1	7,2	7,3	7,8	9,5
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	8,5	8,8	8,1	9,2	9,3	9,8	10,5	10,8
10	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	50	52	51	63	60	62	65	52
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	54	53	65	62	64	67	56	55
11	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	11,1	9,5	9,3	9,5	10,5	10,4	10,5	13,1
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	11,5	11,3	11,5	12,5	12,4	12,5	13,3	13,5
12	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	13,1	11,8	11,8	14,6	13,4	13,1	12	15,1
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	13,8	13,8	16,6	15,4	15,1	14	15,8	18,6
13	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	8,6	9,2	9,5	9,6	8,5	9	9,5	10,6
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	11,2	11,5	11,6	10,5	11	11,5	13,5	13,6
14	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	11,8	10,8	12,3	12,4	12,9	13,1	11,8	13,8
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	12,8	14,3	14,4	14,9	15,1	13,8	16,3	16,4
15	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	29	31	28	31	33	35	38	31
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	33	30	33	35	37	40	32	35

Окончание табл. 4.1

1	2	3	4	5	6	7	8	9	10
16	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	10,4	10,5	12,8	12	11,5	12,6	10,6	12,4
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	12,5	14,8	14	13,5	14,6	12,6	16,8	16
17	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	11,8	11,5	12,2	12,6	12,8	12,9	12,5	13,8
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	13,5	14,2	14,6	14,8	14,9	14,5	16,2	16,6
18	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	9,6	7,3	8,2	10,4	8,8	9,4	8,7	11,6
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	9,3	10,2	12,4	10,8	11,4	10,7	12,2	14,4
19	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	10,5	9,5	9,8	9,1	10,2	10,3	10,8	12,5
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	11,5	11,8	11,1	12,2	12,3	12,8	13,8	13,1
20	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	53	55	54	66	63	65	68	55
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	57	56	68	65	67	70	58	70
21	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	5,1	3,5	3,3	3,5	4,5	4,4	4,5	7,1
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	5,5	5,3	5,5	6,5	6,4	6,5	7,5	7,3
22	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	7,1	5,8	5,8	8,6	7,4	7,1	6	9,1
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	7,8	7,8	10,6	9,4	9,1	8	9,8	9,8
23	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	2,6	3,2	3,5	3,6	2,5	3	3,5	4,6
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	5,2	5,5	5,6	4,5	5	5,5	7,2	7,5
24	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	5,8	4,8	6,3	6,4	6,9	7,1	5,8	7,8
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	6,8	8,3	8,4	8,9	9,1	7,8	8,8	10,3
25	<i>t</i>	1	2	3	4	5	6	7	8
	<i>Y</i>	23	25	22	25	27	29	32	25
	<i>t</i>	9	10	11	12	13	14	15	16
	<i>Y</i>	27	24	27	29	31	34	29	26

4.2. Содержательная постановка задачи и ее математическая модель

Построение эконометрических моделей с целью их дальнейшего применения для прогнозирования предполагает, что оценки, полученные с помощью этих моделей, адекватно отражают действительность. Для этого статистические оценки параметров модели должны иметь такие свойства:

быть **несмещенными**, то есть математическое ожидание оценок параметров модели равна их теоретическим значениям;

быть **состоятельными**, то есть с увеличением объема выборочной совокупности оценка параметра стремится к его значению в генеральной совокупности;

быть **эффективными**, то есть гарантировать наименьшее значение дисперсии по сравнению с оценками, которые получены при расчете с использованием других методов.

Для того чтобы статистические оценки имели такие свойства, необходимо выполнение определенных условий. Эти условия сформулированы в форме **теоремы Гаусса – Маркова**. Согласно этой теореме МНК-оценки параметров модели являются несмещенными, состоятельными и эффективными, если:

- 1) выбрана правильная спецификация модели;
- 2) все внешние факторы являются попарно независимыми;
- 3) ошибки модели являются случайными, то есть математическое ожидание ошибок модели равна нулю;
- 4) дисперсия ошибок является постоянной величиной;
- 5) ошибки модели для различных измерений не коррелируют между собой.

В отличие от пространственных моделей, которые рассматривались в заданиях 1 и 2, последовательные уровни временного ряда не всегда можно считать независимыми. При этом остатки модели временного ряда могут содержать как тенденцию, так и циклические колебания. То есть при построении эконометрических моделей временных рядов необходимо иметь в виду, что достаточно может наблюдаться автокорреляция остатков модели. Нарушение пятого условия теоремы Гаусса – Маркова приводит к тому, что МНК-оценки параметров модели будут несмещенными, но неэффективными, что приводит к увеличению доверительного интервала для прогноза с использованием таких эконометрических моделей.

Итак, в ходе исследования рядов динамики необходимо проверять, являются ли остатки модели независимыми.

Автокорреляция остатков может быть обусловлена несколькими причинами. К ним относятся: наличие ошибок измерения значений эндогенного фактора; отсутствие в спецификации модели экзогенного фактора, влияние которого является существенным; неучет нескольких второстепенных факторов, совместное влияние которых на результативный признак является существенным; неправильный выбор функции, с помощью которой осуществляется аппроксимация.

Исследование наличия автокорреляции остатков модели можно провести с помощью нескольких критериев, среди которых наиболее удобным является **критерий Дарбина – Уотсона**. В этом случае по результатам вычисления ошибок модели, которая была построена по методу МНК, определяют эмпирическое значение критерия:

$$DW = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2}, \quad (4.1)$$

где e_t, e_{t-1} – остатки модели для уровней, которые отвечают последовательным моментам времени t и $t-1$, соответственно;

T – период, для которого осуществляются измерения.

При уровне значимости $\alpha = 0,05$, количестве наблюдений n и количестве объясняющих переменных m определяют нижнюю d_L и верхнюю d_U границы критерия (табл. Д.1 приложения Д) и сравнивают с ними эмпирические значения критерия DW . Правило принятия решения состоит в следующем:

если $0 \leq DW \leq d_L$, то принимают гипотезу о наличии положительной автокорреляции остатков модели;

если $d_L \leq DW \leq d_U$ или $4 - d_U \leq DW \leq 4 - d_L$, то нельзя сделать вывод о наличии или отсутствии автокорреляции;

если $d_U \leq DW \leq 4 - d_U$, автокорреляция считается несущественной;

если $4 - d_L \leq DW \leq 4$, принимается гипотеза о наличии отрицательной автокорреляции остатков модели.

Для описания автокорреляции **первого порядка**, то есть корреляции, которая описывает связь между остатками двух последовательных измерений, применяют уравнение:

$$e_t = \rho \cdot e_{t-1} + \varepsilon_t = \rho(\rho \cdot e_{t-2} + \varepsilon_{t-1}) + \varepsilon_t, \quad (4.2)$$

где ρ – **коэффициент автокорреляции**, отражающий тесноту и направление корреляционной связи между остатками модели текущего и предшествующего уровней, то есть, между двумя рядами остатков, а именно – рядом $e_2; e_3; e_4; \dots; e_t$ и рядом $e_1; e_2; e_3; \dots; e_{t-1}$.

Для приближенного оценивания коэффициента автокорреляции применяют формулу:

$$\rho \approx r = \frac{n}{n-1} \cdot \frac{\sum_{t=2}^T e_t e_{t-1}}{\sum_{t=1}^T e_t^2} + \frac{m}{n}. \quad (4.3)$$

В случае выявления автокорреляции остатков модели, построенной с помощью классического МНК, для отыскания эффективных оценок параметров эконометрической модели применяется **обобщенный метод наименьших квадратов** (ОМНК).

Обобщенный метод наименьших квадратов основан на теореме Эйткена, согласно которой в классе линейных несмещенных оценок неизвестных параметров регрессионной модели наименьшую ковариационную матрицу имеют оценки, полученные как решение матричного уравнения:

$$\widehat{\mathbf{B}} = (\mathbf{X}^T \mathbf{\Omega}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{\Omega}^{-1} \mathbf{Y}, \quad (4.4)$$

где $\widehat{\mathbf{B}}$ – матрица оценок параметров модели;

\mathbf{X} – матрица исходных значений экзогенного фактора, к которой слева приписан столбец единиц;

\mathbf{X}^T – транспонированная матрица \mathbf{X} ;

$\mathbf{\Omega}$ – **ковариационная матрица остатков модели**;

$\mathbf{\Omega}^{-1}$ – матрица, обратная к матрице $\mathbf{\Omega}$;

\mathbf{Y} – матрица исходных значений эндогенного фактора.

Ковариационная матрица остатков модели является квадратной и имеет размер $n \times n$:

$$\Omega = \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \dots & \dots & \dots & \dots & \dots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{pmatrix}, \quad (4.5)$$

При нахождении матрицы Ω^{-1} следует иметь в виду, что при $s > 2$ величина ρ^s стремится к нулю. Отсюда:

$$\Omega^{-1} = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho & 0 & \dots & 0 \\ -\rho & 1+\rho^2 & -\rho & \dots & 0 \\ 0 & -\rho & 1+\rho^2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}. \quad (4.6)$$

Для вычисления произведения матриц и расчета обратных матриц используют определения, приведенные в разделе учебной дисциплины "Линейная алгебра".

4.3. Пример выполнения типового задания № 4

Исходные данные. В табл. 4.2. приведены данные о динамике изменения объемов продаж промышленной продукции Y (млн. грн) в течение семи лет.

Таблица 4.2

Исходные данные типового задания

t , годы	1	2	3	4	5	6	7
Y , млн грн	45	50	48	52	50	59	53

По данным табл. 4.2 построим график временного ряда (рис. 4.1).

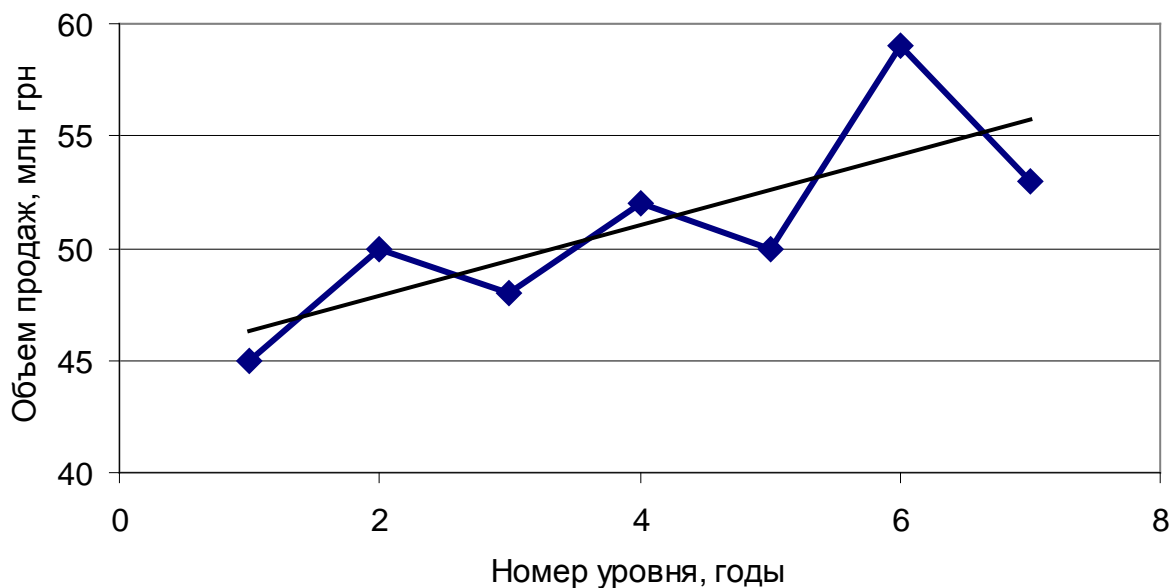


Рис. 4.1. Изменение объема продаж промышленной продукции

Приведенный график (рис. 4.1) свидетельствует о том, что существует тенденция к росту уровней временного ряда (линейный тренд), но при этом отклонения от линии тренда такие, что каждое последующее имеет противоположный знак по сравнению с предыдущим. Это позволяет предположить наличие отрицательной автокорреляции остатков.

Вычислить МНК-оценки параметров уравнения тренда в предположении парной линейной регрессии:

$$\hat{y} = b_0 + b_1 t.$$

Методика вычисления была приведена в данном издании как образец выполнения задания 1. Результаты промежуточных вычислений представлены в табл. 4.3.

Таблица 4.3

Результаты промежуточных вычислений, применяемых при построении эконометрической модели по методу МНК

t	y	t^2	$t \cdot y$	\hat{y}	e_t	e_{t-1}	$(e_t - e_{t-1})^2$	e_t^2	$e_t e_{t-1}$
1	2	3	4	5	6	7	8	9	10
1	45	1	45	46,29	-1,29	–	–	1,65	–
2	50	4	100	47,86	2,14	-1,29	11,76	4,59	-2,76

1	2	3	4	5	6	7	8	9	10
3	48	9	144	49,43	-1,43	2,14	12,76	2,04	-3,06
4	52	16	208	51,00	1,00	-1,43	5,90	1,00	-1,43
5	50	25	250	52,57	-2,57	1,00	12,76	6,61	-2,57
6	59	36	354	54,14	4,86	-2,57	55,18	23,59	-12,49
7	53	49	371	55,71	-2,71	4,86	57,33	7,37	-13,18
28	357	140	1472				155,67	46,86	-35,49
4	51	20	210,29						

Предпоследняя строка табл. 4.3 содержит суммы по столбцам, а последняя – выборочные средние соответствующих характеристик. Эти строки выделены жирным шрифтом.

Вычисляем оценки параметров модели аналогично тому, как это осуществлялось в задании 1. Так, используя значения, приведенные в столбцах 1 – 4 табл. 4.3, получаем:

$$b_1 = \frac{\overline{ty} - \bar{t} \cdot \bar{y}}{\overline{t^2} - \bar{t}^2} = \frac{210,29 - 4 \cdot 51}{20 - 4^2} = 1,57;$$

$$b_0 = \bar{y} - b_1 \cdot \bar{t} = 51 - 1,57 \cdot 4 = 44,71.$$

Теперь записываем уравнение линейного тренда:

$$\hat{y} = 44,71 + 1,57 \cdot t.$$

Проверим наличие (или отсутствие) автокорреляции остатков полученной модели по критерию Дарбина – Уотсона.

Результаты вычислений значение эндогенного фактора по модели линейной регрессии приведены в столбце 5 табл. 4.3, а значения остатков модели, которые вычисляем как разность между значениями, записанными в столбцах 2 и 5, представлены в столбце 6 табл. 4.3. Отсюда находим величины, необходимые для расчета эмпирического значения критерия Дарбина – Уотсона:

$$\sum_{t=2}^T (e_t - e_{t-1})^2 = 155,67; \quad \sum_{t=1}^T e_t^2 = 46,86.$$

Теперь вычислим значение статистики Дарбина – Уотсона:

$$DW = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2} = \frac{155,65}{46,86} = 3,32.$$

При уровне значимости $\alpha = 0,05$ для количества измерений $n = 7$ и количества объясняющих переменных $m = 1$ находим нижнюю границу статистического критерия Дарбина – Уотсона $d_L = 0,70$ и его верхнюю границу $d_U = 1,37$ (см. табл. Д.1 приложения Д). В данном случае выполняется двойное неравенство: $4 - d_L \leq DW \leq 4$, и в соответствии с этим принимаем гипотезу о наличии отрицательной автокорреляции остатков модели, построенной с применением МНК. Отсюда следует, что для обеспечения эффективности оценок параметров эконометрической модели необходимо применение ОМНК.

Определим коэффициент автокорреляции остатков модели. Для этого необходимо определить произведение ошибок $e_t e_{t-1}$ (столбец 10 в табл. 4.3). Отсюда получаем:

$$\rho \approx \frac{n}{n-1} \cdot \frac{\sum_{t=2}^T e_t e_{t-1}}{\sum_{t=1}^T e_t^2} + \frac{m}{n} = \frac{7}{7-1} \cdot \frac{-35,49}{46,86} + \frac{1}{7} = -0,74.$$

Из соотношения (4.6), выполнив несложные вычисления, находим матрицу, обратную к матрице ковариаций остатков модели:

$$\Omega^{-1} = \frac{1}{0,42} \cdot \begin{pmatrix} 1 & 0,74 & 0 & 0 & 0 & 0 & 0 \\ 0,74 & 1,55 & 0,74 & 0 & 0 & 0 & 0 \\ 0 & 0,74 & 1,55 & 0,74 & 0 & 0 & 0 \\ 0 & 0 & 0,74 & 1,55 & 0,74 & 0 & 0 \\ 0 & 0 & 0 & 0,74 & 1,55 & 0,74 & 0 \\ 0 & 0 & 0 & 0 & 0,74 & 1,55 & 0,74 \\ 0 & 0 & 0 & 0 & 0 & 0,74 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 2,22 & 1,64 & 0 & 0 & 0 & 0 & 0 \\ 1,64 & 3,43 & 1,64 & 0 & 0 & 0 & 0 \\ 0 & 1,64 & 3,43 & 1,64 & 0 & 0 & 0 \\ 0 & 0 & 1,64 & 3,43 & 1,64 & 0 & 0 \\ 0 & 0 & 0 & 1,64 & 3,43 & 1,64 & 0 \\ 0 & 0 & 0 & 0 & 1,64 & 3,43 & 1,64 \\ 0 & 0 & 0 & 0 & 0 & 1,74 & 2,22 \end{pmatrix}.$$

Записываем матрицу \mathbf{X}^T как транспонированную матрицу значений экзогенного фактора, к которой слева приписан столбик, состоящий из одних единиц. Соответственно, в транспонированной матрице \mathbf{X}^T значений экзогенного фактора первая строка содержит одни единицы:

$$\mathbf{X}^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{pmatrix}.$$

В соответствии с формулой (4.4) вычислим произведение $\mathbf{X}^T \Omega^{-1}$:

$$\mathbf{X}^T \Omega^{-1} = \begin{pmatrix} 3,86 & 6,72 & 6,72 & 6,72 & 6,72 & 6,72 & 3,86 \\ 5,50 & 13,43 & 20,15 & 26,86 & 33,58 & 40,29 & 25,36 \end{pmatrix}.$$

Теперь вычислим произведение трех матриц:

$$\mathbf{X}^T \Omega^{-1} \mathbf{X} = \begin{pmatrix} 41,29 & 165,17 \\ 165,17 & 787,43 \end{pmatrix}.$$

Вычислим определитель полученной матрицы:

$$\begin{aligned} \Delta \mathbf{X}^T \Omega^{-1} \mathbf{X} &= \begin{vmatrix} 41,29 & 165,17 \\ 165,17 & 787,43 \end{vmatrix} = \\ &= 41,29 \cdot 787,43 - 165,17^2 = 5233,64. \end{aligned}$$

Поскольку определитель матрицы отличен от нуля, то есть матрица является невырожденной, то для нее существует обратная. Находим матрицу, которая является обратной к только что вычисленной:

$$(\mathbf{X}^T \Omega^{-1} \mathbf{X})^{-1} = \begin{pmatrix} 0,150 & -0,032 \\ -0,032 & 0,008 \end{pmatrix}.$$

Итак, мы получили первый из множителей, который содержится в формуле (4.4).

Теперь вычисляем произведение $\mathbf{X}^T \Omega^{-1} \mathbf{Y}$.

Произведение матриц $\mathbf{X}^T \Omega^{-1}$ уже было рассчитано. Получаем:

$$\mathbf{X}^T \Omega^{-1} \mathbf{Y} = \begin{pmatrix} 3,86 & 6,72 & 6,72 & 6,72 & 6,72 & 6,72 & 3,86 \\ 5,50 & 13,43 & 20,15 & 26,86 & 33,58 & 40,29 & 25,36 \end{pmatrix} \cdot \begin{pmatrix} 45 \\ 50 \\ 48 \\ 52 \\ 50 \\ 59 \\ 53 \end{pmatrix} = \begin{pmatrix} 2117,37 \\ 8683,24 \end{pmatrix}.$$

В соответствии с формулой (4.4) определяем ОМНК-оценки параметров модели:

$$\hat{\mathbf{B}} = \begin{pmatrix} 0,150 & -0,032 \\ -0,032 & 0,008 \end{pmatrix} \cdot \begin{pmatrix} 2117,37 \\ 8683,24 \end{pmatrix} = \begin{pmatrix} 44,53 \\ 1,69 \end{pmatrix}.$$

Итак, применив обобщенный метод наименьших квадратов, мы получили модель линейного тренда в следующем виде:

$$\hat{y} = 44,53 + 1,69 \cdot t.$$

Найдем значение эндогенного фактора, используя полученную модель, и проверим по критерию Дарбина – Уотсона, существует ли автокорреляция остатков этой модели. Вычисления приведены в табл. 4.4, последняя строка которой содержит суммы по соответствующим столбцам.

Таблица 4.4

**Результаты вычислений остатков модели
по обобщенному методу наименьших квадратов**

t	\hat{y}	e_t	e_{t-1}	$(e_t - e_{t-1})^2$	e_t^2
1	46,22	-1,22	-	-	1,48
2	47,90	2,10	-1,22	10,98	4,39
3	49,59	-1,59	2,10	13,59	2,53
4	51,28	0,72	-1,59	5,35	0,52
5	52,96	-2,96	0,72	13,59	8,78
6	54,65	4,35	-2,96	53,49	18,92
7	56,34	-3,34	4,35	59,08	11,13
28				156,08	47,76

По результатам вычислений (см. табл. 4.4) находим эмпирическое значение статистики Дарбина – Уотсона, соответствующее эконометрической модели, построенной по методу ОМНК:

$$DW = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2} = \frac{156,08}{47,76} = 3,26.$$

Эмпирическое значение статистического критерия Дарбина – Уотсона, которое получено для эконометрической модели, построенной с использованием обобщенного метода наименьших квадратов, оказалось меньше того, что соответствует модели, построенной по классическому методу наименьших квадратов.

По справочной таблице (см. табл. Д.1 приложения Д) находим, что уровню значимости $\alpha = 0,05$ для количества наблюдений $n = 7$ и количества объясняющих переменных $m = 1$ нижняя граница статистического критерия Дарбина – Уотсона имеет значение $d_L = 0,70$, а верхняя

граница – значение $d_U = 1,37$. Отсюда получаем, что относительно эмпирического значения критерия Дарбина – Уотсона выполняется двойное неравенство: $4 - d_U \leq DW \leq 4 - d_L$. В данном случае нельзя сделать вывод о наличии или отсутствии автокорреляции остатков модели. Таким образом, по сравнению с результатами, которые были получены для модели с использованием классического МНК, эконометрическая модель, полученная с использованием обобщенного метода наименьших квадратов, дает качественно лучший результат. Для иллюстрации различий между результатами, которые получены по методам МНК и ОМНК в приближении линейной аппроксимации, приведем график временного ряда (рис. 4.2).

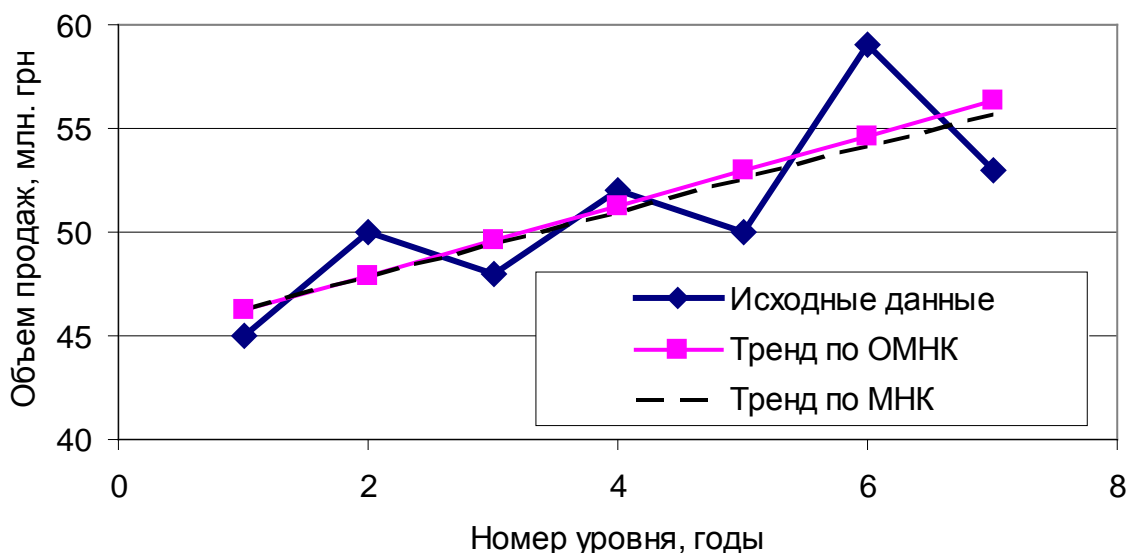


Рис. 4.2. Сравнение вычислений по эконометрическим моделями, которые построены по методам МНК и ОМНК

Вывод. Для анализа динамики изменений объемов продаж промышленной продукции целесообразно применять модель парной регрессии, построенной по методу ОМНК:

$$\hat{y} = 44,53 + 1,69 \cdot t,$$

поскольку согласно теореме Эйткена оценки параметров модели, полученные с помощью обобщенного метода наименьших квадратов, являются эффективными, то есть имеют наименьшую из дисперсий среди оценок, полученных другими методами, а именно – по методу МНК.

Рекомендованная литература

1. Егоршин А. А. Корреляционно-регрессионный анализ : Курс лекций и лабораторных работ : пособ. для вузов / А. А. Егоршин, Л. М. Малярец. – Х. : Основа, 1998. – 208 с.
2. Железнякова Э. Ю. Методические рекомендации к решению индивидуальных заданий по дисциплине "Экономико-математические методы и модели" для иностранных студентов отрасли знаний 0305 "Экономика и предпринимательство" всех форм обучения / Э. Ю. Железнякова, Л. А. Норик. – Х. : ХНЭУ, 2012. – 80 с.
3. Малярец Л. М. Эконометрика в примерах и задачах : учеб. пособ. для иностранных студентов / Л. М. Малярец, Э. Ю. Железнякова, Л. А. Норик. – Х. : Вид. ХНЭУ им. С. Кузнецца, 2014. – 268 с.
4. Малярец Л. М. Экономико-математические методы и модели : учеб. пособ. для иностранных студентов / Л. М. Малярец. – Х. : Изд. ХНЭУ, 2013. – 288 с.
5. Таха Х. А. Введение в исследование операций / Х. А. Таха; пер. с англ. – 7-е изд. – М. : ИД "Вильямс", 2009. – 912 с.
6. Тихомиров Н. П. Эконометрика : учебник / Н. П. Тихомиров, Е. Ю. Дорохина. – М. : Изд. "Экзамен", 2003. – 512 с.
7. Христиановский В. В. Прикладная эконометрия : учебн. для экон. вузов / В. В. Христиановский, Н. Г. Гузь, О. Г. Кривенчуг. – Донецк : ДонГУ, 1998. – 172 с.
8. Экономико-математические методы и модели : учеб. пособ. для вузов / под общ. Ред. А. В. Кузнецова. – 2-е изд. – Минск : БГЭУ, 2010. – 412 с.

Приложения

Приложение А

Таблица А.1

Критические точки распределения F Фишера – Снедекора

(k_1 – количество степеней свободы большей дисперсии,

k_2 – количество степеней свободы меньшей дисперсии)

при уровне значимости $\alpha = 0,05$

k_2	k_1											
	1	2	3	4	5	6	7	8	9	10	11	12
1	161	200	216	225	230	234	237	239	241	242	243	244
2	18,5	19,0	19,2	19,2	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,4
3	10,1	9,55	9,28	9,13	9,01	8,94	8,88	8,84	8,81	8,78	8,76	8,74
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94	2,91
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,31	2,28
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,22	2,18
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	2,12	2,09
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	2,04	2,00
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,98	1,95
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,01	1,97	1,93	1,89
100	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,88	1,85
200	3,89	3,04	2,65	2,41	2,26	2,14	2,05	1,98	1,92	1,87	1,83	1,80
400	3,86	3,02	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,81	1,78
∞	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88	1,83	1,79	1,75

Критические точки распределения F Фишера – Снедекора $(k_1$ – количество степеней свободы большей дисперсии, k_2 – количество степеней свободы меньшей дисперсии)при уровне значимости $\alpha = 0,01$

k_2	k_1								
	1	2	3	4	5	6	7	8	9
1	4052,2	4999,5	5403,3	5624,6	5763,7	5859,0	5928,3	5981,6	6022,5
2	98,503	99,000	99,166	99,249	99,299	99,332	99,356	99,374	99,388
3	34,116	30,817	29,457	28,710	28,237	27,911	27,672	27,489	27,345
4	21,198	18,000	16,694	15,977	15,522	15,207	14,976	14,799	14,659
5	16,258	13,274	12,060	11,392	10,967	10,672	10,456	10,289	10,158
6	13,745	10,925	9,780	9,148	8,746	8,466	8,260	8,102	7,976
7	12,246	9,547	8,451	7,847	7,160	7,191	6,993	6,840	6,719
8	11,259	8,649	7,591	7,006	6,632	6,371	6,178	6,029	5,911
9	10,561	8,022	6,992	6,422	6,057	5,802	5,613	5,467	5,351
10	10,044	7,559	6,552	5,994	5,636	5,386	5,200	5,057	4,942
11	9,646	7,206	6,217	5,668	5,316	5,069	4,865	4,745	4,632
12	9,330	6,927	5,953	5,412	5,064	4,812	4,640	4,499	4,388
13	9,074	6,701	5,739	5,205	4,862	4,620	4,441	4,302	4,191
14	8,862	6,515	5,564	5,035	4,695	4,456	4,278	4,140	4,030
15	8,683	6,359	5,417	4,893	4,556	4,318	4,142	4,005	3,895
16	8,531	6,226	5,292	4,773	4,437	4,202	4,026	3,890	3,780
17	8,400	6,112	5,185	4,669	4,336	4,102	3,927	3,791	3,682
18	8,285	6,013	5,092	4,579	4,248	4,015	3,841	3,705	3,597
19	8,185	5,926	5,010	4,500	4,171	3,939	3,765	3,631	3,523
20	8,096	5,849	4,938	4,431	4,103	3,871	3,699	3,564	3,457
21	8,017	5,780	4,874	4,369	4,042	3,812	3,640	3,506	3,398
22	7,945	5,719	4,817	4,313	3,988	3,758	3,587	3,453	3,346
23	7,881	5,664	4,765	4,264	3,939	3,710	3,539	3,406	3,299
24	7,823	5,614	4,718	4,218	3,895	3,667	3,496	3,363	3,256
25	7,770	5,568	4,676	4,177	3,865	3,627	3,457	3,324	3,217
26	7,721	5,526	4,637	4,140	3,818	3,591	3,421	3,288	3,182
27	7,677	5,488	4,601	4,106	3,785	3,558	3,388	3,256	3,149
29	7,598	5,421	4,538	4,045	3,725	3,500	3,330	3,198	3,092
30	7,568	5,390	4,510	4,018	3,699	3,474	3,305	3,173	3,067
40	7,314	5,179	4,313	3,828	3,514	3,291	3,124	2,993	2,888
60	7,077	4,977	4,126	3,649	3,339	3,119	2,953	2,823	2,719
120	6,851	4,787	3,949	3,480	3,174	2,956	2,792	2,663	2,559
∞	6,635	4,605	3,782	3,319	3,017	2,802	2,639	2,511	2,407

Критические точки распределения Стьюдента

Уровень значимости α (двусторонняя критическая область)				
Число степеней свободы k	$\alpha = 0,01$	$\alpha = 0,02$	$\alpha = 0,05$	$\alpha = 0,1$
1	63,7	31,82	12,7	6,31
2	9,92	6,97	4,30	2,92
3	5,84	4,54	3,18	2,35
4	4,60	3,75	2,78	2,13
5	4,03	3,37	2,57	2,01
6	3,71	3,14	2,45	1,94
7	3,50	3,00	2,36	1,89
8	3,36	2,90	2,31	1,86
9	3,25	2,82	2,26	1,83
10	3,17	2,76	2,23	1,81
11	3,11	2,72	2,20	1,80
12	3,05	2,68	2,18	1,78
13	3,01	2,65	2,16	1,77
14	2,98	2,62	2,14	1,76
15	2,95	2,60	2,13	1,75
16	2,92	2,58	2,12	1,75
17	2,90	2,57	2,11	1,74
18	2,88	2,55	2,10	1,73
19	2,86	2,54	2,09	1,73
20	2,85	2,53	2,09	1,73
21	2,83	2,52	2,08	1,72
22	2,82	2,51	2,07	1,72
23	2,81	2,50	2,07	1,71
24	2,80	2,49	2,06	1,71
25	2,79	2,49	2,06	1,71
26	2,78	2,48	2,06	1,71
27	2,77	2,47	2,05	1,71
28	2,76	2,46	2,05	1,70
30	2,75	2,46	2,04	1,70
40	2,70	2,42	2,02	1,68
60	2,66	2,39	2,00	1,67
120	2,62	2,36	1,98	1,66
∞	2,58	2,33	1,96	1,64
k	$\alpha = 0,005$	$\alpha = 0,01$	$\alpha = 0,025$	$\alpha = 0,05$
Уровень значимости α (односторонняя критическая область)				

Критические значения критерия Дарбина – Уотсона DW при $\alpha = 0,05$

<i>n</i>	Количество независимых переменных											
	<i>m</i> = 1		<i>m</i> = 2		<i>m</i> = 3		<i>m</i> = 4		<i>m</i> = 5		<i>m</i> = 6	
	<i>d_L</i>	<i>d_U</i>	<i>d_L</i>	<i>d_U</i>	<i>d_L</i>	<i>d_U</i>	<i>d_L</i>	<i>d_U</i>	<i>d_L</i>	<i>d_U</i>	<i>d_L</i>	<i>d_U</i>
6	0,61	1,40										
7	0,70	1,37	0,47	1,9								
8	0,76	1,33	0,56	1,78	0,37	2,29						
9	0,82	1,32	0,63	1,70	0,46	2,13	0,30	2,59				
10	0,88	1,32	0,70	1,64	0,53	2,02	0,38	2,41	0,24	2,82		
11	0,93	1,32	0,76	1,60	0,60	1,93	0,44	2,28	0,32	2,65	0,12	2,89
12	0,97	1,33	0,81	1,58	0,66	1,86	0,51	2,18	0,38	2,51	0,16	2,67
13	1,01	1,34	0,86	1,56	0,72	1,82	0,57	2,09	0,45	2,39	0,21	2,49
14	1,05	1,35	0,91	1,55	0,77	1,78	0,63	2,03	0,51	2,30	0,26	2,35
15	1,08	1,36	0,95	1,54	0,81	1,75	0,69	1,98	0,56	2,22	0,30	2,24
16	1,11	1,37	0,98	1,54	0,86	1,73	0,73	1,94	0,62	2,16	0,35	2,15
18	1,16	1,39	1,05	1,54	0,93	1,70	0,82	1,87	0,71	2,06	0,44	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,89	1,83	0,79	1,99	0,52	1,92
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94	1,57	1,85
25	1,29	1,45	1,21	1,55	1,12	1,65	1,04	1,77	0,95	1,89	1,68	1,78
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85	0,76	1,73
32	1,37	1,50	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82	0,86	1,69
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80	0,91	1,67
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72	1,23	1,79	1,00	1,64
45	1,48	1,57	1,43	1,62	1,38	1,67	1,34	1,72	1,29	1,78	1,07	1,64
50	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72	1,34	1,77	1,12	1,64
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73	1,41	1,77	1,21	1,64
70	1,58	1,64	1,55	1,67	1,53	1,70	1,49	1,74	1,46	1,77	1,28	1,65
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77	1,34	1,65
90	1,64	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78	1,38	1,66
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78	1,42	1,67
150	1,72	1,75	1,71	1,76	1,69	1,77	1,68	1,79	1,67	1,80	1,54	1,71
200	1,76	1,78	1,75	1,79	1,74	1,80	1,73	1,81	1,72	1,82	1,61	1,74

Содержание

Введение	3
Задание 1. Эконометрическая модель парной регрессии.....	5
1.1. Задания для самостоятельного выполнения	5
1.2. Содержательная постановка задачи и ее математическая модель	8
1.3. Пример выполнения типового задания № 1.....	12
Задание 2. Эконометрическая модель многофакторной регрессии	22
2.1. Задания для самостоятельного выполнения	22
2.2. Содержательная постановка задачи и ее математическая модель.....	25
2.3. Пример выполнения типового задания № 2.....	30
Задание 3. Моделирование временного ряда: анализ и прогноз	38
3.1. Задания для самостоятельного выполнения	38
3.2. Содержательная постановка задачи и ее математическая модель.....	41
3.3. Пример выполнения типового задания № 3.....	44
Задание 4. Исследование автокорреляции остатков временного ряда	55
4.1. Задания для самостоятельного выполнения	55
4.2. Содержательная постановка задачи и ее математическая модель.....	58
4.3. Пример выполнения типового задания № 4.....	61
Рекомендованная литература.....	69
Приложения.....	70

НАВЧАЛЬНЕ ВИДАННЯ

**Методичні рекомендації
до виконання контрольних робіт
з навчальної дисципліни
"ЕКОНОМЕТРИКА"
для іноземних студентів
заочної форми навчання
(рос. мовою)**

Самостійне електронне текстове мережеве видання

Укладачі: **Железнякова** Еліна Юріївна
Лебедєва Ірина Леонідівна
Норік Лариса Олексіївна

Відповідальний за видання *Л. М. Малярець*

Редактор *Н. І. Ганцевич*

Коректор *Н. І. Ганцевич*

У збірнику вміщено задачі до контрольних робіт з навчальної дисципліни та приклади виконання нульового варіанта цих робіт. Розв'язання задач супроводжується коротким викладом теоретичних положень, що допоможе студентам заочної форми навчання отримати теоретичні знання та набути практичних навичок у їх використанні.

Рекомендовано для іноземних студентів заочної форми навчання.

План 2016 р. Поз. 285-ЕВ. Обсяг 75 с.

Видавець і виготовлювач – ХНЕУ ім. С. Кузнеця, 61166, м. Харків, просп. Науки, 9-А
*Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру
ДК № 4853 від 20.02.2015 р.*