Федеральное государственное образовательное бюджетное учреждение
высшего образования
«ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
Новороссийский филиал
Кафедра «Информатики, математики и общегуманитарные науки»

И.Г. РЗУН

Методические рекомендации

ЭКОНОМЕТРИЧЕСКИЕ ИССЛЕДОВАНИЯ

Направление подготовки: 38.04.01 Экономика
Направленность программы : Учёт и корпоративные финансы
Форма обучения: очная, заочная
Квалификация (степень) выпускника: магистр

Новороссийск 2018

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ИЗУЧЕНИЮ ДИСЦИПЛИНЫ

Цели и задачи

Целью изучения дисциплины является формирование у обучающихся системного представления об эконометрике, как науке, исследующей данные статистики для изучения поведения, описания и прогнозирования развития различных факторов.

Задачи дисциплины

Важной методической задачей курса является формирование понимания обучающимися основных положений эконометрики; приобретение опыта построения эконометрических моделей, принятия решений о спецификации и идентификации модели и выбора метода оценки параметров модели, интерпретации результатов, получения прогнозных оценок на основе анализа эконометрических данных; освоение современных эконометрических пакетов прикладных программ.

РЕКОМЕНДАЦИИ ПО ПРОВЕДЕНИЮ ЛЕКЦИЙ

При подготовке к занятиям обучающийся должен просмотреть конспекты лекций, рекомендованную литературу по данной теме; подготовиться к ответу на контрольные вопросы. Успешное изучение дисциплины требует от обучающихся посещения лекций, активной работы на семинарах, выполнения всех учебных заданий преподавателя, ознакомления с основной и дополнительной литературой, интернет-источниками.

Запись лекции – одна из форм активной самостоятельной работы обучающихся, требующая навыков и умения кратко, схематично, последовательно и логично фиксировать основные положения, выводы, обобщения, формулировки. Культура записи лекции – один из важнейших факторов успешного и творческого овладения знаниями. Последующая работа над текстом лекции воскрешает в памяти содержание лекции, позволяет развивать аналитическое мышление. Лекции имеют обзорный характер и нацелены на освещение наиболее трудных и дискуссионных вопросов, а также призваны способствовать формированию навыков самостоятельной работы с научной литературой. Работа с конспектом лекций предполагает просмотр конспекта в тот же день после занятий, пометку материала конспекта, который вызывает затруднения для понимания. Попытайтесь найти ответы на затруднительные вопросы, используя рекомендуемую литературу. Если самостоятельно не удалось разобраться в материале, сформулируйте вопросы и обратитесь за помощью к преподавателю на консультации, ближайшей лекции или семинаре. Регулярно отводите время для повторения пройденного материала, проверяя свои знания, умения и навыки по контрольным вопросам.

Работу с основной и дополнительной литературой целесообразно начинать с освоения материала учебников, которые содержат необходимый материал по каждой теме.

Подготовка к семинарскому занятию зависит от темы занятия и вопросов, предложенных преподавателем, для подготовки к семинару.

Постоянная активность на занятиях, готовность ставить и обсуждать актуальные проблемы дисциплины – залог успешной работы и положительной оценки.

ВВЕДЕНИЕ

Моделирование экономических процессов сопряжено с рядом трудностей. Это и многообразие экономической жизни и конфликт интересов различных социальных групп и внешний фактор в силу открытости современной экономики. Возникает определенный пессимизм по отношению к возможностям и полезности количественного моделирования, стремление к качественному описанию взаимосвязей экономических величин. Тем не менее, конкретные решения, влекущие материальную ответственность, не могут опираться на качественные рассуждения и требуют точных вычислений. Востребованные практикой средства анализа данных, на которые можно опираться в процессе принятия решений, предоставляет эконометрика. В этой науке соединились возможности экономической теории и математики.

В предлагаемом пособии представлены контрольные задания по пяти основным темам, соответствующим программе по эконометрике. В первом задании предлагается построить линейную регрессионную модель с одним фактором, влияющим на результат (парная регрессия). В рамках построенной модели требуется получить оценки параметров, оценить качество модели. Во втором задании необходимо построить нелинейную модель с теми же исходными данными. В третьем задании число регрессоров (факторов) увеличивается до двух (множественная регрессия). Необходимо оценить параметры, проверить нулевые гипотезы относительно значений параметров, оценить качество множественной регрессии, сделать прогноз. В четвертом задании предлагается идентифицировать параметры системы одновременных уравнений. В пятом задании изучается автокорреляционная функция временного ряда, и прогнозирование в условиях авторегрессии.

По каждому из пяти заданий предлагаются методические указания, которые включают теоретические выкладки и примеры решения эконометрических задач. Уровень сложности предлагаемых заданий и относительно небольшое количество наблюдений позволяют выполнить предлагаемую работу с помощью обычного калькулятора. Допускается использование специализированных пакетов программ, например оболочки EXCEL.

Выполнение работы следует начинать с проработки методических указаний, параллельно изучая теорию в соответствии со стандартом и рабочей программой курса. Затем выполняются задания своего варианта. При подготовке к экзамену, рекомендуется письменно ответить на вопросы для самопроверки.

1. ТРЕБОВАНИЯ К ОФОРМЛЕНИЮ КОНТРОЛЬНОЙ РАБОТЫ. ВЫБОР ВАРИАНТА

Задания выполняются с применением компьютера. Вычисления производятся с точностью до трех знаков после запятой. Номер варианта N определяется по двум последним цифрам nn Вашей зачетной книжки по правилу:

$N=20$, если $nn=00$, $N=nn$, если $00<nn\le 20$; $N=nn-20$, если $20<nn\le 40$; $N=nn-40$, если $40<nn\le 60$, и т.д.

Объем контрольной работы зависит от решения задачи. Оформление должно полностью соответствовать стандартам, разработанным в университете. Первый лист контрольной работы – титульный лист (это лист с номером 1, на нем номер листа не проставляется). Остальные листы, включая приложение, нумеруются сквозной нумерацией.

Выполненное задание контрольной работы сдается в двух файлах: один – текстовый документ, с титульным листом и подробным описанием этапов проектирования базы данных; второй – файл созданной базы данных.

Контрольная работа должна быть правильно оформлена. Текст контрольной работы набирается на компьютере с использованием текстового редактора MS Word:

– страницы должны иметь сквозную нумерацию;

– каждую структурную часть работы следует начинать с нового листа;

– необходимо стремиться к ясности, краткости и самостоятельности изложения материала;

– в тексте работы не должно быть сокращений слов, за исключением общепринятых;

– при представлении табличного материала над правым верхним углом таблицы помещают надпись «Таблица» с указанием ее порядкового номера (например, «Таблица 5»), снабжают тематическим заголовком, который располагают посередине страницы и пишут с прописной буквы без точки в конце;

– приводимые в работе иллюстрации (схема, диаграмма, график, технический рисунок, фотография) должны быть выполнены четко, аккуратно, разборчиво и иметь номер и подрисуночную подпись (например, «Рисунок 4.Вид отчёта»);

– гарнитура шрифта — Times New Roman;

– размер шрифта — 14;

– межстрочный интервал — полуторный;

– размеры полей: левого — 3 см, верхнего — 2, правого — 1,5, нижнего — 2 см;

– ориентация — книжная, при необходимости-альбомная;

– форматирование основного текста и ссылок — по ширине;

– цвет шрифта — авто;

– абзацный отступ — 1,25 см.

В окончательном виде оформите работу одним файлом в формате doc (через команду «**Сохранить как»** выберите соответствующий тип файла и в конец имени файла через нижнее подчеркивание добавьте учебный год, группу и фамилию с инициалами.

2. КОНТРОЛЬНЫЕ ЗАДАНИЯ

Задание 1

Линейная регрессионная модель

В табл.1 даны наблюдения x_t и y_t . Предполагается, что зависимую переменную y и независимую x связывает линейное регрессионное уравнение $y_t=a+b\cdot x_t+\varepsilon_t$,
где a и b неизвестные параметры уравнения, ε_t – случайные отклонения.

1. Постройте диаграмму рассеяния наблюдений и визуально проверьте гипотезу о возможной линейной зависимости между x и y;
2. По методу наименьших квадратов (МНК) определите оценки параметров a и b линейной регрессионной модели;
3. На диаграмме рассеяния постройте график прогнозных значений $\hat{y}_t = \hat{a} + \hat{b} \cdot x_t$, где $\hat{a}$ - оценка параметра a, а $\hat{b}$ - оценка параметра b.
4. Вычислите оценку дисперсии остатков. Оцените дисперсию $\hat{a}$ и $\hat{b}$;
5. С уровнем значимости 0,05 проверьте гипотезу a=100 и гипотезу b=0.
6. Постройте 95% доверительные интервалы для параметров a и b.
7. Определите коэффициент детерминации R^2, качественно оценить тесноту связи между x и y;
8. Вычислите дисперсионное отношение F, с уровнем значимости 0,05 проверьте гипотезу о наличии связи между x и y;

9. Определите прогнозное значение $\hat{y}_{11}$ при x_{11}=N, где N – номер Вашего варианта. Постройте 95% доверительный интервал для найденного прогнозного значения.
10. Оцените с помощью эластичности силу влияния фактора на результат в точке x_{11}.

Задание 2

Нелинейная модель. Линеаризация

Для тех же наблюдений x_t и y_t, предполагается, что зависимую переменную y и независимую x связывает нелинейное регрессионное уравнение:

$y_t = a + b \cdot \frac{1}{x_t} + \varepsilon_t$ для вариантов 1,6,11,16;

$y_t = a \cdot b^{x_t} \cdot \varepsilon_t$ для вариантов 2,7,12,17;

$\ln y_t = a + b \cdot \frac{1}{x_t} + \varepsilon_t$ для вариантов 3,8,13,18;

$y_t = a \cdot e^{b \cdot x_t} \cdot \varepsilon_t$ для вариантов 4,9,14,19;

$y_t = a \cdot x_t{}^b \cdot \varepsilon_t$ для вариантов 5,10,15,20.

1. Проведите линеаризацию модели, определите оценки параметров нелинейной модели.
2. Оцените качество модели с помощью коэффициента детерминации и дисперсионного отношения F.
3. Определите прогнозное значение $\hat{y}_{11}$ при x_{11}=N, где N – номер Вашего варианта. Постройте 95% доверительный интервал для прогноза.

4. Оцените с помощью эластичности силу влияния фактора на результат в точке x_{11}.
5. На диаграмме рассеяния постройте график прогнозных значений. Определите сумму квадратов отклонений наблюдений от нелинейного прогноза.

Таблица 1

t	Вариант 1		Вариант 2		Вариант 3		Вариант 4		Вариант 5	
	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t
1	9	31	15	60	13	25	14	110	19	26
2	11	31	17	84	9	22	18	136	11	33
3	15	35	7	19	15	25	16	125	6	42
4	6	28	17	75	15	25	8	84	9	36
5	17	33	15	59	18	27	19	140	15	28
6	13	31	16	65	6	18	6	77	17	28
7	14	34	14	50	11	24	16	120	18	28
8	16	34	13	55	14	25	13	100	11	34
9	7	28	6	32	18	27	8	84	19	25
10	11	31	15	53	8	21	6	75	11	33
t	Вариант 6		Вариант 7		Вариант 8		Вариант 9		Вариант 10	
	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t
1	9	131	7	16	9	25	12	97	10	33
2	7	130	11	26	18	31	6	74	19	25
3	16	144	14	38	20	31	14	107	14	31
4	11	140	9	25	17	29	13	108	17	28
5	17	157	7	21	6	17	18	128	19	27

6	8	123	15	40	7	23	14	113	14	29
7	19	155	10	24	7	21	20	141	6	43
8	18	159	12	32	8	21	14	109	16	28
9	12	145	6	18	20	30	10	88	9	39
10	14	153	7	19	18	30	14	112	10	36
t	Вариант 11		Вариант 12		Вариант 13		Вариант 14		Вариант 15	
	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t
1	16	32	13	52	9	22	5	72	19	26
2	19	34	15	66	8	20	18	136	19	25
3	12	31	15	62	20	27	10	88	13	30
4	16	32	14	61	9	22	7	75	19	27
5	14	34	6	26	15	26	15	114	7	41
6	6	29	19	89	15	26	10	97	17	26
7	13	32	11	41	14	26	5	73	9	37
8	10	32	14	54	19	26	8	84	18	27
9	12	34	7	25	18	27	19	143	8	38
10	20	35	16	70	19	27	7	79	9	37

t	Вариант 16		Вариант 17		Вариант 18		Вариант 19		Вариант 20	
	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t
1	17	158	7	23	9	24	8	84	19	27
2	15	138	6	18	11	28	19	140	16	27
3	8	129	14	36	17	32	6	77	19	26
4	5	114	7	21	19	27	16	120	14	29
5	16	155	16	39	14	32	13	100	17	29
6	14	151	19	52	9	27	8	84	14	32

7	20	163	14	37	19	33	6	75	12	31
8	16	145	9	20	16	31	12	97	6	45
9	12	144	16	41	12	27	6	74	14	30
10	12	148	6	17	6	21	14	107	19	26

Задание 3
Множественная регрессия

К тем же наблюдениям x_t, и y_t. добавляются значения $z_t = \sqrt{x_t}$. Предполагается, что зависимую переменную y и факторы связывает уравнение множественной линейной регрессии

$$y_t = a + b \cdot x_t + c \cdot z_t + \varepsilon_t,$$

где a, b и c неизвестные параметры уравнения, ε_t – случайные отклонения.

1. Определите МНК оценки параметров уравнения.
2. С уровнем значимости 0,05 проверьте гипотезу b=0 (о влиянии фактора x на результат) и c=0 (о влиянии фактора z на результат).
3. Определите коэффициент детерминации и скорректированный коэффициент детерминации.
4. По критерию Фишера F с уровнем значимости 0,05 оцените качество модели в целом.
5. Составьте корреляционную таблицу наблюдений и вычислите частные коэффициенты корреляции.
6. Сравните по качеству модели заданий 1, 2 и 3.

Задание 4

Системы регрессионных уравнений

В каждом из заданий (табл.2) предлагается структурная система эконометрических уравнений, приведенная система уравнений и данные наблюдений.

1. Определите к какому типу относится каждое из уравнений структурной системы эконометрических уравнений (идентифицируемо, неиндентифицируемо или сверхидентифицируемо).
2. Опираясь на данные наблюдений и построенную на их основе приведенную систему эконометрических уравнений, проведите идентификацию параметров структурной системы.

Таблица 2

Вариант 1.

Структурная система:

$$y_{t1} = b_{10} + b_{12} y_{t2} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21} y_{t1} + b_{23} y_{t3} + a_{23} x_{t3} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31} y_{t1} + a_{31} x_{t1} + a_{32} x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 2 - x_{t1} + 3x_{t2} - x_{t3}$$
$$\hat{y}_{t2} = -1 + 2x_{t1} - 2x_{t2} + x_{t3}$$
$$\hat{y}_{t3} = 5 + 2x_{t1} - 4x_{t2} + 3x_{t3}$$

Данные наблюдений:

t	y_t	y_t	y_t	x_t	x_t	x_t

Вариант 2.

Структурная система:

$$y_{t1} = b_{10} + b_{12} y_{t2} + b_{13} y_{t3} + a_{13} x_{t3} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21} y_{t1} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31} y_{t1} + a_{31} x_{t1} + a_{32} x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 5 + 3x_{t1} + 2x_{t2} + 5x_{t3}$$
$$\hat{y}_{t2} = 1 + 4x_{t1} - 2x_{t2} - 2x_{t3}$$
$$\hat{y}_{t3} = -1 - 3x_{t1} + x_{t2} + x_{t3}$$

Данные наблюдений:

t	y_t	y_t	y_t	x_t	x_t	x_t

	1	2	3	1	2	3
1	14	-8	-5	1	5	3
2	6	0	3	2	3	2
3	20	-7	-8	4	8	3
4	-5	14	23	6	2	5
5	6	2	14	2	4	7

	1	2	3	1	2	3
1	26	-5	-1	1	1	3
2	19	4	-4	2	2	1
3	47	-4	-1	4	7	3
4	32	19	-16	6	3	1
5	48	-5	-1	3	4	5

Вариант 3.

Структурная система:

$$y_{t1} = b_{10} + b_{12} y_{t2} + a_{12} x_{t2} + a_{13} x_{t3} + \varepsilon_{t1}$$

$$y_{t2} = b_{20} + b_{21} y_{t1} + b_{23} y_{t3} + a_{21} x_{t1} + \varepsilon_{t2}$$

$$y_{t3} = b_{30} + b_{32} y_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 2 - 3x_{t1} + x_{t2} + 4x_{t3}$$

$$\hat{y}_{t2} - 4 + 2x_{t1} - x_{t2} + 2x_{t3}$$

$$\hat{y}_{t3} = -3 - 2x_{t1} + 2x_{t2} + 5x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	-6	3	-1	1	5	3
2	15	0	20	2	3	2
3	7	-1	1	4	8	3

Вариант 4.

Структурная система:

$$y_{t1} = b_{10} + b_{12} y_{t2} + \varepsilon_{t1}$$

$$y_{t2} = b_{20} + b_{21} y_{t1} + a_{21} x_{t1} + a_{23} x_{t3} + \varepsilon_{t2}$$

$$y_{t3} = b_{30} + b_{31} y_{t1} + b_{32} y_{t2} + a_{32} x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = -3 + 4x_{t1} + x_{t2} - 5x_{t3}$$

$$\hat{y}_{t2} = 1 - 2x_{t1} - 2x_{t2} + 3x_{t3}$$

$$\hat{y}_{t3} = 5 - 1x_{t1} - x_{t2} - x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	11	-12	-9	1	1	3
2	-1	-6	-4	2	2	1

			7			
4	4	18	13	6	2	5
5	29	20	38	2	4	7

3	3	-13	-7	4	7	3
4	-5	4	-8	6	3	1
5	-26	12	-7	3	4	5

Вариант 5.

Структурная система:

$$y_{t1} = b_{10} + b_{12} y_{t2} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21} y_{t1} + b_{23} y_{t3} + a_{23} x_{t3} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31} y_{t1} + a_{31} x_{t1} + a_{32} x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 1 + x_{t1} + 2x_{t2} + 2x_{t3}$$
$$\hat{y}_{t2} = -1 + 3x_{t1} + x_{t2} + 3x_{t3}$$
$$\hat{y}_{t3} = 2 + 3x_{t1} + 2x_{t2} + 3x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	26	35	44	1	5	3
2	13	17	21	2	3	2
3	14	16	25	4	8	3
4	13	23	24	6	2	5

Вариант 6.

Структурная система:

$$y_{t1} = b_{10} + b_{12} y_{t2} + b_{13} y_{t3} + a_{13} x_{t3} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21} y_{t1} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31} y_{t1} + a_{31} x_{t1} + a_{32} x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 1 + 3x_{t1} + x_{t2} - 2x_{t3}$$
$$\hat{y}_{t2} = -2 + 2x_{t1} + 3x_{t2} + 2x_{t3}$$
$$\hat{y}_{t3} = 1 + 3x_{t1} + 4x_{t2} + 5x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	13	25	47	1	1	3
2	15	18	33	2	2	1
3	16	35	62	4	7	3
4	11	27	50	6	3	1

5	20	17	26	2	4	7

Вариант 7.

Структурная система:

$$y_{t1} = b_{10} + b_{12}y_{t2} + a_{12}x_{t2} + a_{13}x_{t3} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21}y_{t1} + b_{23}y_{t3} + a_{21}x_{t1} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{32}y_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 1 + 2x_{t1} + 2x_{t2} + 2x_{t3}$$
$$\hat{y}_{t2} = 2 - 1x_{t1} - 2x_{t2} + x_{t3}$$
$$\hat{y}_{t3} = -1 - 4x_{t1} + 4x_{t2} + 3x_{t3}$$

Данные наблюдений:

t	y_t	y_t	y_t	x_t	x_t	x_t

5	10	16	28	3	4	5

Вариант 8.

Структурная система:

$$y_{t1} = b_{10} + b_{12}y_{t2} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21}y_{t1} + a_{21}x_{t1} + a_{23}x_{t3} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31}y_{t1} + b_{32}y_{t2} + a_{32}x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 1 + 1x_{t1} + 4x_{t2} + 2x_{t3}$$
$$\hat{y}_{t2} = -2 - 1x_{t1} + 1x_{t2} - 3x_{t3}$$
$$\hat{y}_{t3} = 2 + 2x_{t1} + x_{t2} + 2x_{t3}$$

Данные наблюдений:

t	y_t	y_t	y_t	x_t	x_t	x_t

	1	2	3	1	2	3
1	26	-11	11	1	5	3
2	30	-3	-18	2	3	2
3	14	-4	15	4	8	3
4	29	-9	37	6	2	5
5	16	-4	9	2	4	7

	1	2	3	1	2	3
1	30	-10	16	1	1	3
2	24	-12	25	2	2	1
3	44	-15	29	4	7	3
4	23	-19	25	6	3	1
5	16	-15	15	3	4	5

Вариант 9.

Структурная система:

$$y_{t1} = b_{10} + b_{12} y_{t2} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21} y_{t1} + b_{23} y_{t3} + a_{23} x_{t3} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31} y_{t1} + a_{31} x_{t1} + a_{32} x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 2 + 2x_{t1} + x_{t2} + 2x_{t3}$$
$$\hat{y}_{t2} = 3 - 3x_{t1} - 2x_{t2} + x_{t3}$$
$$\hat{y}_{t3} = 1 + x_{t1} + x_{t2} + 3x_{t3}$$

Данные наблюдений:

t	y_t	y_t	y_t	x_t	x_t	x_t

Вариант 10.

Структурная система:

$$y_{t1} = b_{10} + b_{12} y_{t2} + b_{13} y_{t3} + a_{13} x_{t3} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21} y_{t1} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31} y_{t1} + a_{31} x_{t1} + a_{32} x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 2 + x_{t1} + 2x_{t2} + 2x_{t3}$$
$$\hat{y}_{t2} = 6 - 4x_{t1} - 5x_{t2} + 2x_{t3}$$
$$\hat{y}_{t3} = -2 + 3x_{t1} + 4x_{t2} + x_{t3}$$

Данные наблюдений:

t	y_t	y_t	y_t	x_t	x_t	x_t

	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	16	-9	15	1	5	3
2	12	-6	12	2	3	2
3	25	-23	24	4	8	3
4	24	-12	23	6	2	5
5	25	-4	28	2	4	7

	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	12	1	7	1	1	3
2	9	-9	13	2	2	1
3	27	-40	43	4	7	3
4	14	-29	28	6	3	1
5	24	-16	28	3	4	5

Вариант 11.

Структурная система:

$$y_{t1} = b_{10} + b_{12}y_{t2} + a_{12}x_{t2} + a_{13}x_{t3} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21}y_{t1} + b_{23}y_{t3} + a_{21}x_{t1} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{32}y_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 6 - 3x_{t1} + x_{t2} + 4x_{t3}$$
$$\hat{y}_{t2} = 2 + x_{t1} - x_{t2} + 2x_{t3}$$
$$\hat{y}_{t3} = -3 - 2x_{t1} + 2x_{t2} + 3x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	-2	3	-3	1	5	3

Вариант 12.

Структурная система:

$$y_{t1} = b_{10} + b_{12}y_{t2} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21}y_{t1} + a_{21}x_{t1} + a_{23}x_{t3} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31}y_{t1} + b_{32}y_{t2} + a_{32}x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 2 + 4x_{t1} + x_{t2} - 3x_{t3}$$
$$\hat{y}_{t2} = -3 + 2x_{t1} - 2x_{t2} - 2x_{t3}$$
$$\hat{y}_{t3} = 5 - 1x_{t1} + 4x_{t2} - x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	22	-7	11	1	1	3

2	19	5	14	2	3	2
3	11	1	13	4	8	3
4	8	18	3	6	2	5
5	33	22	20	2	4	7

2	8	-12	21	2	2	1
3	14	-20	33	4	7	3
4	10	-1	-3	6	3	1
5	-7	-19	8	3	4	5

Вариант 13.

Структурная система:

$$y_{t1} = b_{10} + b_{12} y_{t2} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21} y_{t1} + b_{23} y_{t3} + a_{23} x_{t3} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31} y_{t1} + a_{31} x_{t1} + a_{32} x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 2 + x_{t1} + 3x_{t2} - 4x_{t3}$$
$$\hat{y}_{t2} = -4 + 5x_{t1} - 3x_{t2} + 3x_{t3}$$
$$\hat{y}_{t3} = 3 + x_{t1} + 4x_{t2} + 3x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	14	28	39	1	5	3

Вариант 14.

Структурная система:

$$y_{t1} = b_{10} + b_{12} y_{t2} + b_{13} y_{t3} + a_{13} x_{t3} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21} y_{t1} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31} y_{t1} + a_{31} x_{t1} + a_{32} x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 1 + 3x_{t1} + 5x_{t2} - 5x_{t3}$$
$$\hat{y}_{t2} = 2 + 3x_{t1} + 2x_{t2} + 2x_{t3}$$
$$\hat{y}_{t3} = 1 - 5x_{t1} - 4x_{t2} + 5x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	24	28	-25	1	1	3

2	-8	8	24	2	3	2
3	12	9	24	4	8	3
4	-15	20	23	6	2	5
5	15	-6	35	2	4	7

Вариант 15.

Структурная система:

$$y_{t1} = b_{10} + b_{12}y_{t2} + a_{12}x_{t2} + a_{13}x_{t3} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21}y_{t1} + b_{23}y_{t3} + a_{21}x_{t1} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{32}y_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 1 + 2x_{t1} - 3x_{t2} - 4x_{t3}$$
$$\hat{y}_{t2} = 1 - 2x_{t1} - 2x_{t2} + x_{t3}$$
$$\hat{y}_{t3} = -2 - 4x_{t1} + 2x_{t2} + 3x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	-17	-16	0	1	5	3
2	-5	-1	-2	2	3	2

2	13	27	-23	2	2	1
3	39	35	-34	4	7	3
4	0	37	-14	6	3	1
5	23	18	-20	3	4	5

Вариант 16.

Структурная система:

$$y_{t1} = b_{10} + b_{12}y_{t2} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21}y_{t1} + a_{21}x_{t1} + a_{23}x_{t3} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31}y_{t1} + b_{32}y_{t2} + a_{32}x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 3 + x_{t1} + 3x_{t2} + x_{t3}$$
$$\hat{y}_{t2} = -3 - x_{t1} - x_{t2} + 2x_{t3}$$
$$\hat{y}_{t3} = 1 - x_{t1} + 4x_{t2} + 2x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	24	-6	24	1	1	3
2	21	-9	9	2	2	1
3	3	-5	3	4	7	3

		3	1			
3	-13	-6	8	4	8	3
4	-41	-12	20	6	2	5
5	-11	-7	2	2	4	7

	4		7			
4	18	1	12	6	3	1
5	13	2	11	3	4	5

Вариант 17.

Структурная система:

$$y_{t1} = b_{10} + b_{12}y_{t2} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21}y_{t1} + b_{23}y_{t3} + a_{23}x_{t3} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31}y_{t1} + a_{31}x_{t1} + a_{32}x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = -1 + 2x_{t1} + 4x_{t2} + 2x_{t3}$$
$$\hat{y}_{t2} = 3 - 3x_{t1} - 2x_{t2} + 4x_{t3}$$
$$\hat{y}_{t3} = -2 + x_{t1} + 3x_{t2} + 3x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	28	0	22	1	5	3
2	18	0	15	2	3	2
3	46	-14	37	4	8	3

Вариант 18.

Структурная система:

$$y_{t1} = b_{10} + b_{12}y_{t2} + b_{13}y_{t3} + a_{13}x_{t3} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21}y_{t1} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31}y_{t1} + a_{31}x_{t1} + a_{32}x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 1 + x_{t1} + 2x_{t2} + 4x_{t3}$$
$$\hat{y}_{t2} = 5 + 4x_{t1} + 6x_{t2} + 2x_{t3}$$
$$\hat{y}_{t3} = -4 - 3x_{t1} - 5x_{t2} + x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	17	19	-10	1	1	3
2	10	28	-19	2	2	1
3	3	6	-	4	7	3

4	27	3	24	6	2	5
5	34	17	33	2	4	7

	2	8	46			
4	15	51	-37	6	3	1
5	33	51	-28	3	4	5

Вариант 19.

Структурная система:

$$y_{t1} = b_{10} + b_{12}y_{t2} + a_{12}x_{t2} + a_{13}x_{t3} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21}y_{t1} + b_{23}y_{t3} + a_{21}x_{t1} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{32}y_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 1 - 3x_{t1} + 3x_{t2} + x_{t3}$$
$$\hat{y}_{t2} = -3 - 2x_{t1} - x_{t2} + 2x_{t3}$$
$$\hat{y}_{t3} = 1 - 2x_{t1} + 2x_{t2} + x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	0	-20	-1	1	5	3
2	15	-3	12	2	3	2
3	16	-16	13	4	8	3
4	-8	-5	-3	6	2	5

Вариант 20.

Структурная система:

$$y_{t1} = b_{10} + b_{12}y_{t2} + \varepsilon_{t1}$$
$$y_{t2} = b_{20} + b_{21}y_{t1} + a_{21}x_{t1} + a_{23}x_{t3} + \varepsilon_{t2}$$
$$y_{t3} = b_{30} + b_{31}y_{t1} + b_{32}y_{t2} + a_{32}x_{t2} + \varepsilon_{t3}.$$

Приведенная система:

$$\hat{y}_{t1} = 3 + 4x_{t1} - 2x_{t2} - 5x_{t3}$$
$$\hat{y}_{t2} = -2 - 2x_{t1} + 3x_{t2} + 3x_{t3}$$
$$\hat{y}_{t3} = 5 - x_{t1} + 4x_{t2} - x_{t3}$$

Данные наблюдений:

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	5	5	11	1	1	3
2	-10	16	21	2	2	1
3	-15	24	33	4	7	3
4	-2	6	-3	6	3	1

5	5	5	6	2	4	7	5	-29	24	8	3	4	5

Задание 5
Временные ряды. Авторегрессия

В табл.3 представлены наблюдения временного ряда.

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	20
	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t
1	50	52	49	49	52	49	52	50	54	52	100	100	98	100	98
2	50	50	49	51	55	47	51	51	55	46	100	99	101	101	101
3	52	52	49	48	47	48	52	52	55	44	101	100	103	103	102
4	51	48	51	50	48	49	51	51	52	51	104	102	98	99	99
5	51	51	49	50	55	49	51	50	53	48	104	100	102	101	101
6	51	49	50	52	53	50	52	49	52	50	103	102	102	102	95
7	52	52	51	50	56	49	51	51	53	44	103	101	99	104	100
8	51	51	52	53	50	49	51	49	49	41	103	103	101	102	101
9	52	52	51	51	50	49	49	49	52	47	103	101	101	103	98
10	51	52	50	55	58	50	50	49	52	46	103	104	100	104	96
11	51	53	51	51	56	49	46	46	51	47	104	101	102	106	93
12	52	54	52	53	59	51	48	49	48	42	103	104	102	105	99
13	52	53	51	53	53	50	48	49	52	40	102	103	102	104	97
14	54	54	53	58	54	49	49	46	50	45	102	104	103	106	96
15	52	53	52	54	60	48	48	48	52	43	101	105	103	108	94
16	53	54	51	56	58	50	48	48	47	45	102	103	104	107	91
17	53	54	54	55	62	48	48	46	50	41	102	106	105	107	96
18	51	54	54	59	56	47	46	46	47	38	104	104	104	108	95
19	51	55	52	56	57	48	46	47	49	41	104	104	105	110	94
20	50	55	55	57	62	48	45	46	45	43	103	103	105	108	92
21	50	55	54	58	60	48	46	45	47	40	103	103	105	108	89
22	51	56	53	60	65	49	45	47	46	37	102	102	106	110	90
23	54	55	55	57	59	47	46	45	46	36	102	104	108	111	91
24	54	57	57	59	60	48	46	45	44	38	103	106	107	111	93
25	53	55	57	61	64	47	47	46	45	38	104	105	107	110	90
26	53	56	55	62	62	47	46	45	44	36	104	106	108	111	87
27	53	56	57	59	67	45	46	45	44	33	107	105	108	112	87

28	53	55	57	59	63	46	46	44	41	34	105	106	108	111	89
29	53	55	56	63	62	45	46	43	42	34	106	106	107	111	90
30	54	55	59	62	68	46	45	44	42	35	104	106	108	112	87

1. Постройте диаграмму наблюдений временного ряда. Определите для него линейный тренд. Вычислите отклонения наблюдений от тренда (остатки регрессии). Установите, является ли данный тренд значимым.
2. Определите и постройте выборочную автокорреляционную функцию остатков (r_i для $i=1,2,..,5$). Установите пиковое значение автокорреляционной функции. Постройте соответствующую найденному пиковому значению модель временного ряда с корреляцией остатков. Оцените качество построенной модели.
3. С помощью построенной модели сделайте прогноз для следующих за тридцатым пяти наблюдений временного ряда.

3. МЕТОДИЧЕСКИЕ УКАЗАНИЯ

3.1. Линейная регрессионная модель

Для анализа работы торгового предприятия произведено 10 наблюдений числа покупателей x_t и выручки y_t (табл.4):

Таблица 4

t	1	2	3	4	5	6	7	8	9	10
x_t	31	75	89	26	35	73	91	21	56	21
y_t	64	100	103	50	63	95	109	43	93	37

Предполагается, что зависимую переменную (выручку) и независимую (число покупателей в магазине) связывает линейное регрессионное уравнение $y_t = a + bx_t + \varepsilon_t$.

1. Построим диаграмму рассеяния наблюдений (рис.1), откладывая на координатной плоскости 10 точек с координатами (31; 64), (75; 100),…, (21; 37):

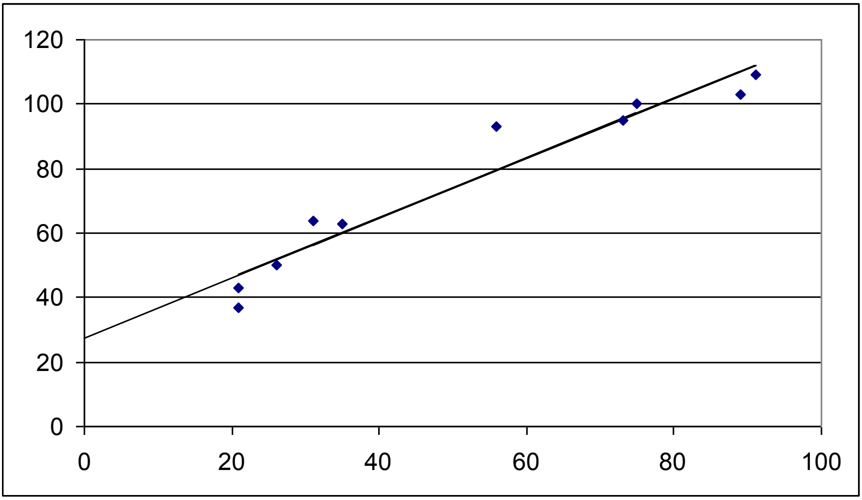

Рис.1 Диаграмма рассеяния наблюдений (точки), линейный тренд (сплошная прямая).

По типу диаграммы рассеяния можно предположить, что между наблюдениями x и y существует линейная зависимость.

2. Применяя метод наименьших квадратов, получим оценки параметров a и b линейной регрессионной модели:

Оценка параметра b вычисляется по формуле

$$\hat{b} = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{4589{,}8 - 51{,}8\cdot 75{,}7}{3403{,}6 - 51{,}8^2} = 0{,}928,$$

где $\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t = 51{,}8$; $\bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t = 75{,}7$; $\overline{x^2} = \frac{1}{n}\sum_{t=1}^{n} x_t^2 = 3403{,}6$;

$\overline{xy} = \frac{1}{n}\sum_{t=1}^{n} x_t y_t = 4589{,}8$;

n – число наблюдений. В представленном примере n=10.

Оценка параметра a вычисляется по формуле

$$\hat{a} = \bar{y} - \hat{b}\cdot\bar{x} = 75{,}7 - 0{,}928\cdot 51{,}8 = 27{,}626.$$

Для получения оценок параметров модели удобно использовать табл.5:

Таблица 5

t	x_t	y_t	x_t^2	y_t^2	$x_t y_t$	$\hat{y}_t$	e_t	e_t^2
1	31	64	961	4096	1984	56,396	7,604	57,817
2	75	100	5625	10000	7500	97,231	2,769	7,667
3	89	103	7921	10609	9167	110,224	-7,224	52,186
4	26	50	676	2500	1300	51,756	-1,756	3,083
5	35	63	1225	3969	2205	60,109	2,891	8,361
6	73	95	5329	9025	6935	95,375	-0,375	0,141
7	91	10	8281	118	9919	112,0	-	9,487

		9		81		80	3,080	
8	21	43	441	1849	903	47,116	-4,116	16,938
9	56	93	3136	8649	5208	79,598	13,402	179,617
10	21	37	441	1369	777	47,116	-10,116	102,326
Σ	518	757	34036	63947	45898	757,000	0,000	437,623
Σ/n	51,8	75,7	3403,6	6395	4589,8	75,700	0,000	

Можно сделать следующие выводы:

- среднее число покупателей 51,8;
- средняя выручка 75,7 ед.;
- каждый покупатель приносит в среднем 0,928 ед. выручки.

3. Уравнение прогнозных значений имеет вид:

$$\hat{y}_t = \hat{a} + \hat{b} \cdot x_t = 27{,}626 + 0{,}928 \cdot x_t .$$

Заполним соответствующий столбец в таблице и построим график прогнозных значений на диаграмме рассеяния.

4. Остатки линейной регрессионной модели определим по формуле

$$e_t = y_t - \hat{y}_t$$

Оценка дисперсии остатков равна

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{n} e_t^2}{n-2} = \frac{437{,}623}{10-2} = 54{,}703.$$

Оценка дисперсии $\hat{a}$ равна

$$\hat{\sigma}_{\hat{a}}^2 = \frac{\hat{\sigma}^2\,\overline{\overline{x^2}}}{n\left(\overline{\overline{x^2}} - \bar{x}^2\right)} = 25{,}846\ .$$

Оценка дисперсии $\hat{b}$ равна

$$\hat{\sigma}_{\hat{b}}^2 = \frac{\hat{\sigma}^2}{n\left(\overline{\overline{x^2}} - \bar{x}^2\right)} = 0{,}00759\ .$$

5. Гипотеза $a = a_0$ будет проверяться исходя из того, что случайная величина

$$t = \frac{\hat{a} - a_0}{\hat{\sigma}_{\hat{a}}}$$

в нормальной классической линейной регрессионной модели подчиняется распределению Стьюдента с (n-2) степенями свободы. Если $|t|$ окажется меньше некоторого критического значения t_α, которое находится по таблицам критических точек распределения Стьюдента, то гипотеза принимается. Если больше, то гипотеза отвергается. Таблица критических точек распределения Стьюдента приводится в Приложении 1.

По таблице находим, что для уровня значимости 0,05 и восьми степеней свободы

t_α=2,306.

Проверим нулевую гипотезу H_0: a=0, при конкурирующей H_1: $a \neq 0$. Вычислим

$$t = \frac{27{,}626}{\sqrt{25{,}846}} = 5{,}434.$$

Поскольку 5,434>2,306, то нулевая гипотеза отвергается.

Выше сказанное справедливо и для параметра b. Проверим гипотезу b=1, которая означает, что один покупатель в среднем приносит торговой точке единицу выручки. Вычислим

$$\left| \frac{0{,}928 - 1}{\sqrt{0{,}00759}} \right| = 0{,}826.$$

Поскольку 0,826<2,306, то гипотеза принимается.

6. Из неравенств

$$\left| \frac{\hat{a} - a}{\hat{\sigma}_{\hat{a}}} \right| < t_\alpha \text{ и } \left| \frac{\hat{b} - b}{\hat{\sigma}_{\hat{b}}} \right| < t_\alpha,$$

находим

$$\hat{a} - t_\alpha \hat{\sigma}_{\hat{a}} < a < \hat{a} + t_a \hat{\sigma}_{\hat{a}}$$

и

$$\hat{b} - t_\alpha \hat{\sigma}_{\hat{b}} < b < \hat{b} + t_\alpha \hat{\sigma}_{\hat{b}}.$$

Подставляя найденные ранее значения, находим 95% доверительные интервалы

$$15{,}903 < a < 39{,}350 \text{ и } 0{,}727 < b < 1{,}129.$$

Последнее неравенство означает, что с вероятностью 0,95 средняя выручка, которую приносит один покупатель, принадлежит интервалу (0,727; 1,129).

7. Коэффициент детерминации равен отношению суммы квадратов отклонений регрессии к общей сумме квадратов отклонений:

$$R^2 = \frac{\sum_{t=1}^{n}(\hat{y}_t - \bar{y})^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2}.$$

Можно доказать, что для парной регрессии данное отношение равно

$$R^2 = \hat{b}^2 \frac{S_x^2}{S_y^2} = 0{,}934,$$

где $S_x^2 = \overline{x^2} - \bar{x}^2 = 720{,}36$ и $S_y^2 = \overline{y^2} - \bar{y}^2 = 664{,}21$ выборочные дисперсии.

В построенной модели дисперсия результата на 93,4% объясняется линейной зависимостью выручки от числа покупателей и только на 6,6% дисперсией неучтенных факторов. Полученное значение коэффициента детерминации близко к единице. Поэтому связь между x и y сильная (число покупателей заметно влияет на выручку данного предприятия).

8. Если фактор (в нашей задаче число покупателей) не влияет на результат (выручку), тогда дисперсионное отношение

$$F = (n-2)\frac{R^2}{1-R^2} = 113{,}422$$

в классической нормальной линейной регрессионной модели, подчиняется распределению Фишера с (1; n-2) числом степеней свободы. По таблице критических точек распределения Фишера (Приложение 2) находим, что для уровня значимости α=0,05, величина F_α=10,128 (число степеней свободы: 1; 8). Найденное значение $F \gg F_\alpha$, что указывает на сильное влияние фактора на результат.

9. Допустим, что планируется расширение предприятия, при этом среднее количество покупателей должно вырасти на 20% и составит $x_{11} = 51{,}8 \cdot 1{,}2 = 67$ чел. Необходимо определить среднюю выручку, которую в этом случае получит предприятие. Выручку найдем из прогнозного уравнения

$$\hat{y}_{11} = \hat{a} + \hat{b} \cdot x_{11} = 27{,}626 + 0{,}928 \cdot 67 = 89{,}807\,.$$

Определим, насколько точным является данный прогноз. Для этого построим 95% доверительный интервал для прогнозируемой выручки:

$$\hat{y}_{11} - \hat{\sigma}_{e_{11}} \cdot t_\alpha < y_{11} < \hat{y}_{11} + \hat{\sigma}_{e_{11}} \cdot t_\alpha$$

$$71{,}66 < y_{11} < 107{,}953,$$

где $\hat{\sigma}_{e_{11}} = \sigma\sqrt{1 + \dfrac{1}{n} + \dfrac{(x_{11} - \bar{x})^2}{nS_x^2}} = 7{,}869$ - стандартное отклонение e_{11}.

10. Эластичность в произвольной точке определяется по формуле

$$\varepsilon_{x_t} = \hat{y}_t' \frac{x_t}{\hat{y}_t},$$

где $\hat{y}_t' = (\hat{a} + \hat{b}x_t)' = \hat{b}$ - производная по фактору x. В точке x_{11} эластичность равна

$$\varepsilon_{x_{11}} = \hat{b}\frac{x_{11}}{\hat{a} + \hat{b}x_{11}} = 0{,}692 .$$

Найденная величина означает, что при увеличении числа покупателей на 1% выручка возрастает в среднем на 0,692%. Справедливо это в окрестности точки x_{11}.

3.2. Нелинейная модель. Линеаризация

Построим нелинейную регрессионную модель в виде

$$y_t = a \cdot x_t^b \cdot \varepsilon_t .$$

Тип нелинейной зависимости определяется формой диаграммы рассеяния, содержанием и теоретической моделью соответствующего экономического процесса. Иногда приходится применять различные нелинейные модели, а затем выбирать из них лучшую.

1. Предложенная модель становится линейной после логарифмирования:

$$\ln y_t = \ln a + b \ln x_t + \ln \varepsilon_t .$$

Обозначая: $Y_t = \ln y_t$; $A = \ln a$; $X_t = \ln x_t$; $u_t = \ln \varepsilon_t$, получим

$$Y_t = A + bX_t + u_t .$$

Преобразуем таблицу исходных данных и, опираясь на результаты вычислений в табл.6, вычислим оценки

$$\hat{A} = 1{,}779; \quad \hat{a} = e^{\hat{A}} = 5{,}923; \hat{b} = 0{,}654 .$$

Уравнение прогнозных значений имеет вид

$$\widehat{Y}_t = 1{,}779 + 0{,}654 \cdot X_t .$$
$$\qquad (0{,}199) \quad (0{,}052)$$

В уравнении в скобках показаны стандартные отклонения оценок параметров.

Прогнозирование в нелинейной модели может осуществляться по формуле

$$\widehat{y}_t = 5{,}923 \cdot x_t^{0{,}654}$$

или с помощью экспоненты

$$\widehat{y}_t = e^{\widehat{Y}_t} .$$

t	$X_t = \ln x_t$	$Y_t = \ln y_t$	X_t^2	Y_t^2	X_tY_t	$\widehat{Y}_t$	$\widehat{y}_t = e^{\widehat{Y}_t}$
1	3,434	4,159	11,792	17,296	14,282	4,023	55,878
2	4,317	4,605	18,641	21,208	19,883	4,601	99,546
3	4,489	4,635	20,148	21,481	20,804	4,712	111,328
4	3,258	3,912	10,615	15,304	12,746	3,908	49,810
5	3,555	4,143	12,640	17,166	14,730	4,102	60,491
6	4,290	4,554	18,408	20,738	19,538	4,583	97,803
7	4,511	4,691	20,348	22,009	21,162	4,727	112,956
8	3,045	3,761	9,269	14,147	11,451	3,769	43,321
9	4,025	4,533	16,203	20,544	18,245	4,410	82,243
10	3,045	3,611	9,269	13,039	10,994	3,769	43,321
Σ	38	42,604	147,334	182,930	163,834	42,604	756,697
	3,797	4,260	14,733	18,293	16,383	4,260	75,670

2. Вычислим коэффициент детерминации

$$R^2 = \widehat{b}^2 \frac{S_X^2}{S_Y^2} = 0{,}654 \frac{0{,}317}{0{,}142} = 0{,}952 ,$$

где $S_X^2 = \overline{X^2} - \overline{X}^2$ и $S_Y^2 = \overline{Y^2} - \overline{Y}^2$.

Дисперсионное отношение

$$F = (n-2)\frac{R^2}{1-R^2} = 158{,}484.$$

Найденное значение F>> F_α=10,128, что указывает на сильное влияние фактора на результат в логарифмической модели.

Чтобы сравнить по качеству нелинейную и линейную модель, вычислим сумму квадратов отклонений нелинейного прогноза от наблюдений

$$\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 = 321{,}112,$$

и определим отношение, которое называют псевдодетерминацией:

$$\hat{R}^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2} = 1 - \frac{321{,}112}{6642{,}1} = 0{,}952.$$

Сравнивая значение 0,952 с коэффициентом детерминации линейной модели 0,934, делаем вывод о том, что нелинейная модель лучше согласуется с данными наблюдений.

3. Как и в предыдущем задании, считаем, что планируется расширение предприятия, при этом среднее количество покупателей должно вырасти на 20% и составит $x_{11} = 51{,}8 \cdot 1{,}2 = 67$ чел. Прогноз средней выручки в нелинейной модели

$$\hat{y}_{11} = 5{,}923 \cdot 67^{0{,}654} = 92{,}471.$$

Определим, насколько точным является данный прогноз. Для этого построим 95% доверительный интервал для логарифма прогнозируемой выручки:

$$\hat{Y}_{11} - \hat{\sigma}_{e_{11}} \cdot t_\alpha < Y_{11} < \hat{Y}_{11} + \hat{\sigma}_{e_{11}} \cdot t_\alpha$$

$$4{,}298 < Y_{11} < 4{,}756,$$

где $\hat{\sigma}_{e_{11}} = \sigma\sqrt{1 + \frac{1}{n} + \frac{(X_{11} - \overline{X})^2}{nS_X^2}} = 0{,}099$ - стандартное отклонение e_{11};

X_{11}=ln x_{11}=ln 67=4,205; $\hat{Y}_{11} = 1{,}779 + 0{,}654 \cdot 4{,}205 = 4{,}527$.

Для прогнозируемой выручки в рамках нелинейной модели доверительный интервал имеет вид

$$e^{4{,}298} < y_{11} < e^{4{,}756}$$

$$73{,}564 < y_{11} < 116{,}238.$$

4. Для степенной нелинейной модели эластичность в произвольной точке определяется по формуле

$$\varepsilon_{x_t} = \hat{y}_t{}' \frac{x_t}{y_t} = \hat{a} \cdot \hat{b} \cdot x_t^{\hat{b}-1} \frac{x_t}{\hat{a} \cdot x_t^{\hat{b}}} = \hat{b} = 0{,}654.$$

Найденная величина означает, что при увеличении среднего числа покупателей на 1% выручка возрастает в среднем на 0,654%.

5.

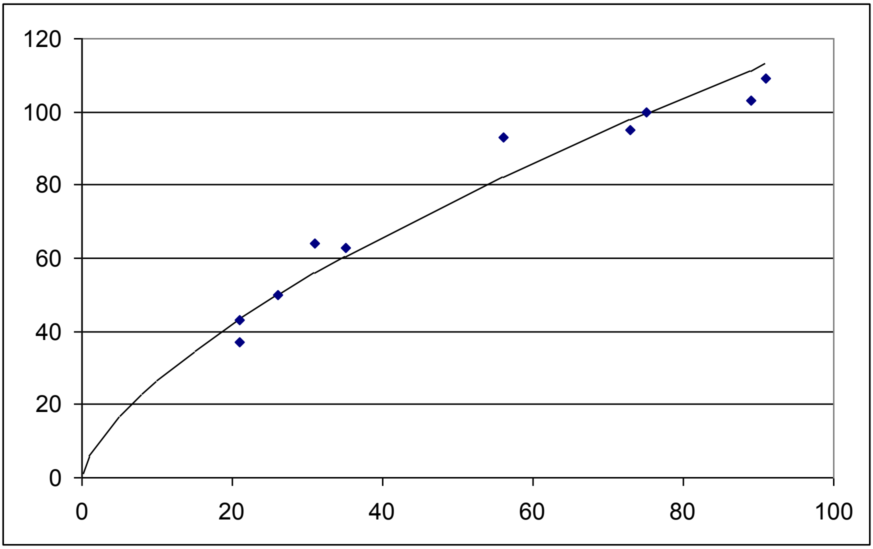

Рис.2. Диаграмма рассеяния наблюдений (точки), нелинейный тренд (сплошная линия).

3.3. Множественная регрессия

На выручку в торговом предприятии могло повлиять наличие или отсутствие рекламы. Чтобы определить эффективность рекламных мероприятий введем фиктивную переменную z_t, которая принимает значение 1, когда реклама применялась, и 0, когда рекламы не было. Данные наблюдений представлены в табл.7.

Таблица 7

t	1	2	3	4	5	6	7	8	9	10
x_t	31	75	89	26	35	73	91	21	56	21
z_t	0	1	1	0	0	0	1	1	1	0
y_t	64	100	103	50	63	95	109	43	93	37

1. Линейное уравнение множественной регрессии при наличии двух факторов, влияющих на результат, имеет вид

$$y_t = a + b \cdot x_t + c \cdot z_t + \varepsilon_t .$$

В соответствии с МНК, оценки параметров множественной регрессии являются решением системы нормальных уравнений:

$$\begin{cases} \hat{a} + \hat{b} \cdot \overline{x} + \hat{c} \cdot \overline{z} = \overline{y} \\ \hat{a} \cdot \overline{x} + \hat{b} \cdot \overline{x^2} + \hat{c} \cdot \overline{xz} = \overline{xy} \\ \hat{a} \cdot \overline{z} + \hat{b} \cdot \overline{xz} + \hat{c} \cdot \overline{z^2} = \overline{zy} \end{cases}$$

Коэффициенты при неизвестных в данной системе уравнений найдем с помощью табл.8:

Таблица 8

t	x_t	z_t	y_t	x_t^2	z_t^2	y_t^2	$x_t z_t$	$x_t y_t$	$y_t z_t$	$\hat{y}_t$	e_t	e_t^2
1	31	0	64	961	0	4096	0	1984	0	56,109	7,891	62,275
2	75	1	100	5625	1	10000	75	7500	100	97,495	2,505	6,277
3	89	1	103	7921	1	10609	89	9167	103	110,346	-7,346	53,969
4	26	0	50	676	0	2500	0	1300	0	51,519	-1,519	2,306
5	35	0	63	1225	0	3969	0	2205	0	59,780	3,220	10,366
6	73	0	95	5329	0	9025	0	6935	0	94,664	0,336	0,113
7	91	1	109	8281	1	11881	91	9919	109	112,182	-3,182	10,127
8	21	1	43	441	1	1849	21	903	43	47,924	-4,924	24,242
9	56	1	93	3136	1	8649	56	5208	93	80,053	12,947	167,625
10	21	0	37	441	0	1369	0	777	0	46,929	-9,929	98,579
Σ	518	5	757	34036	5	63947	332	45898	448	757	0,0	435,880

Σ/n	51,8	0,5	75,7	3403,6	0,5	6394,7	33,2	4589,8	44,8	75,7	0,0	

С найденными коэффициентами система нормальных уравнений приобретает вид

$$\begin{cases} \hat{a} + \hat{b} \cdot 51{,}8 + \hat{c} \cdot 0{,}5 = 75{,}7 \\ \hat{a} \cdot 51{,}8 + \hat{b} \cdot 3403{,}6 + \hat{c} \cdot 33{,}2 = 4589{,}8 \\ \hat{a} \cdot 0{,}5 + \hat{b} \cdot 33{,}2 + \hat{c} \cdot 0{,}5 = 44{,}8. \end{cases}$$

Полученная система уравнений может быть решена методом Крамера.

Найдем определитель матрицы коэффициентов при неизвестных

$$\Delta = \begin{vmatrix} 1 & 51{,}8 & 0{,}5 \\ 51{,}8 & 3403{,}6 & 33{,}2 \\ 0{,}5 & 33{,}2 & 0{,}5 \end{vmatrix} = 126{,}8.$$

Заменим первый столбец свободными коэффициентами и вычислим определитель

$$\Delta_1 = \begin{vmatrix} 75{,}7 & 1{,}8 & 0{,}5 \\ 4589{,}8 & 3403{,}6 & 33{,}2 \\ 44{,}8 & 33{,}2 & 0{,}5 \end{vmatrix} = 3506{,}160.$$

Аналогично заменим второй и третий столбцы свободными коэффициентами и вычислим определители

$$\Delta_2 = \begin{vmatrix} 1 & 75{,}7 & 0{,}5 \\ 51{,}8 & 4589{,}8 & 33{,}2 \\ 0{,}5 & 44{,}8 & 0{,}5 \end{vmatrix} = 116{,}4$$

и

$$\Delta_3 = \begin{vmatrix} 1 & 51,8 & 75,7 \\ 51,8 & 3403,6 & 4589,8 \\ 0,5 & 33,2 & 44,8 \end{vmatrix} = 126,16.$$

В соответствии с методом Крамера оценки параметров равны:

$$\hat{a} = \frac{\Delta_1}{\Delta} = \frac{3506,16}{126,8} = 27,651;$$

$$\hat{b} = \frac{\Delta_2}{\Delta} = \frac{116,4}{126,8} = 0,918;$$

$$\hat{c} = \frac{\Delta_3}{\Delta} = \frac{126,16}{126,8} = 0,995.$$

Дисперсия остатков оценивается по формуле

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{n} e_t^2}{n-k-1} = \frac{435,880}{10-2-1} = 62,269\,.$$

Оценки дисперсий параметров и их стандартные отклонения равны:

$$\hat{\sigma}^2_{\hat{a}} = \hat{\sigma}^2 \cdot \frac{A_{11}}{n \cdot \Delta} = 62,269 \cdot \frac{599,56}{10 \cdot 126,8} = 29,443\,; \qquad \hat{\sigma}_{\hat{a}} = 5,426$$

;

$$\sigma^2_{\hat{b}} = \sigma^2 \cdot \frac{A_{22}}{n \cdot \Delta} = 0,0123\,; \qquad \hat{\sigma}_{\hat{b}} = 0,111$$

;

$$\sigma^2_{\hat{c}} = \hat{\sigma}^2 \cdot \frac{A_{33}}{n \cdot \Delta} = 35,375\,; \qquad \sigma^2_{\hat{c}} = 5,948$$

,

где

$$A_{11} = \begin{vmatrix} \overline{x^2} & \overline{xz} \\ \overline{xz} & \overline{z^2} \end{vmatrix} = \begin{vmatrix} 3403,6 & 33,2 \\ 33,2 & 0,5 \end{vmatrix} = 599,56;$$

$$A_{22} = \begin{vmatrix} 1 & \overline{z} \\ \overline{z} & \overline{z^2} \end{vmatrix} = \begin{vmatrix} 1 & 0,5 \\ 0,5 & 0,5 \end{vmatrix} = 0,25;$$

$$A_{33} = \begin{vmatrix} 1 & \overline{x} \\ \overline{x} & \overline{x^2} \end{vmatrix} = \begin{vmatrix} 1 & 51,8 \\ 51,8 & 3403,6 \end{vmatrix} = 720,36$$ -

соответствующие алгебраические дополнения.

.

Прогнозное уравнение имеет вид

$$\hat{y}_t = \hat{a} + \hat{b} \cdot x_t + \hat{c} \cdot z_t = \underset{(5,426)}{27,651} + \underset{(0,111)}{0,918} \cdot x_t + \underset{(5,948)}{0,995} \cdot z_t.$$

С помощью данного уравнения заполним в таблице столбцы $\hat{y}_t$, $e_t = y_t - \hat{y}_t$ и e_t^2.

2. Если $b = b_0$, то в нормальной классической линейной регрессионной модели случайная величина

$$t = \frac{\hat{b} - b_0}{\hat{\sigma}_{\hat{b}}}$$

подчиняется распределению Стьюдента с n-k-1=10-2-1=7 степенями свободы. Как и в парной регрессии, если $|t| < t_\alpha$, то принимается гипотеза $b = b_0$, в противном случае эта гипотеза отвергается.

Проверяем гипотезу b=0, т.е. b_0=0. Находим отношение

$$|t| = \left|\frac{\hat{b}}{\hat{\sigma}_{\hat{b}}}\right| = \frac{0,918}{0,111} = 8,270\,.$$

По таблице Приложения 1 находим $t_{\alpha} = 2,365$. Поскольку 8,270>2,365, то гипотеза b=0 отвергается, фактор x значимо влияет на результат.

Проверяем гипотезу c=0. Находим отношение

$$|t| = \left|\frac{\hat{c}}{\hat{\sigma}_{\hat{c}}}\right| = \frac{0,995}{5,948} = 0,167\,.$$

Поскольку 0,167<2,365, то гипотеза c=0 принимается, фактор z не влияет на результат. В нашем случае нет оснований считать, что рекламная компания повлияла на выручку. Следует учитывать, что десяти наблюдений для получения значимых результатов в множественной регрессии чаще всего оказывается недостаточно. Как правило, их требуется на порядок больше.

3. Коэффициент детерминации определяется по формуле

$$R^2 = \frac{RSS}{TSS},$$

где $RSS = \sum_{t=1}^{n}(\hat{y}_t - \bar{y})^2$ - сумма квадратов регрессии;

$$TSS = \sum_{t=1}^{n}(y_t - \bar{y})^2 = n(\overline{y^2} - \bar{y}^2) = 10 \cdot (6394,7 - 75,7^2) = 6642,10 -$$

общая сумма квадратов. Для МНК оценок параметров линейной регрессии справедливо

TSS=RSS+ESS,

где $ESS = \sum_{t=1}^{n}(y_t - \hat{y}_t)^2 = \sum_{t=1}^{n} e_t^2 = 435{,}88$ - сумма квадратов ошибок.

Найдем RSS=TSS-ESS=6642,10-435,88=6206,22. Коэффициент детерминации равен

$$R^2 = \frac{6206{,}22}{6642{,}10} = 0{,}934.$$

Скорректированный коэффициент детерминации равен

$$\tilde{R}^2 = 1 - \frac{n-1}{n-3}(1 - R^2) = 1 - \frac{9}{7}(1 - 0{,}934) = 0{,}916.$$

Сравнивая значение скорректированного коэффициента детерминации с коэффициентом детерминации, полученным в задании 1, делаем вывод, что введение дополнительного регрессора z не улучшило качества регрессионной модели.

4. Дисперсионное отношение Фишера равно

$$F = \frac{n-k-1}{k} \cdot \frac{RSS}{ESS} = \frac{10-2-1}{2} \cdot \frac{6206{,}22}{435{,}88} = 49{,}834,$$

где k – число регрессоров в модели (в нашем случае x и z всего два регрессора). С помощью таблицы Приложения 2 находим $F_\alpha = 4{,}737$, число степеней свободы 2 и 7. Поскольку 49,834>4,737 модель является значимой.

5. Выборочный коэффициент корреляции между x и y равен

$$r_{xy} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{S_x S_y} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{\overline{x^2} - \bar{x}^2}\sqrt{\overline{y^2} - \bar{y}^2}} = \frac{4589{,}8 - 51{,}8 \cdot 75{,}7}{\sqrt{3403{,}6 - 51{,}8^2}\sqrt{6394{,}7 - 75{,}7^2}} = 0{,}9$$

.

Аналогично определяются коэффициенты корреляции x и z, а также между z и y. Результаты расчетов представлены в табл.9.

Таблица 9

Коэффициенты корреляции			
r	y	x	z
y	1	0,966	0,539
x	0,966	1	0,544
z	0,539	0,544	1

В общем случае корреляционная матрица наблюдений имеет вид $r = \begin{pmatrix} 1 & r_{yx} & r_{yz} \\ r_{yx} & 1 & r_{xz} \\ r_{yz} & r_{xz} & 1 \end{pmatrix}$.

Алгебраические дополнения ее элементов определяются стандартно. Например

$$A_{12} = -\begin{vmatrix} r_{yx} & r_{xz} \\ r_{yz} & 1 \end{vmatrix} = -(r_{yx} - r_{xz} r_{yz}) .$$

Частный коэффициент корреляции между x и y при фиксированном значении z равен

$$r_{yx.z} = -\frac{A_{12}}{\sqrt{A_{11}}\sqrt{A_{22}}} = \frac{r_{yx} - r_{xz}r_{yz}}{\sqrt{1-r_{xz}^2}\sqrt{1-r_{yz}^2}} = \frac{0{,}966 - 0{,}544 \cdot 0{,}539}{\sqrt{1-0{,}544^2}\sqrt{1-0{,}539^2}} = 0{,}953$$

.

Частный коэффициент корреляции между x и z при фиксированном значении y равен

$$r_{xz.y} = -\frac{A_{23}}{\sqrt{A_{22}}\sqrt{A_{33}}} = \frac{r_{xz} - r_{yx}r_{yz}}{\sqrt{1-r_{yx}^2}\sqrt{1-r_{yz}^2}} = \frac{0{,}544 - 0{,}966 \cdot 0{,}539}{\sqrt{1-0{,}966^2}\sqrt{1-0{,}539^2}} = 0{,}105$$

.

Частный коэффициент корреляции между y и z при фиксированном значении x равен

$$r_{yz.x} = -\frac{A_{13}}{\sqrt{A_{22}}\sqrt{A_{33}}} = \frac{r_{yz} - r_{xz}r_{yx}}{\sqrt{1-r_{xz}^2}\sqrt{1-r_{yx}^2}} = \frac{0{,}539 - 0{,}544 \cdot 0{,}966}{\sqrt{1-0{,}544^2}\sqrt{1-0{,}966^2}} = 0{,}063$$

.

Результаты расчетов частных коэффициентов корреляции представлены в табл.10.

Таблица 10

Коэффициенты частной корреляции			
r	y	x	z
y	1	0,953	0,063
x	0,953	1	0,105
z	0,063	0,105	1

3.4. Системы регрессионных уравнений

Задана структурная система эконометрических уравнений:

$$\begin{cases} y_{t1} = b_{10} + b_{12}y_{t2} + a_{11}x_{t1} + a_{12}x_{t2} + \varepsilon_{t1} \\ y_{t2} = b_{20} + b_{21}y_{t1} + b_{23}y_{t3} + a_{23}x_{t3} + \varepsilon_{t2} \\ y_{t3} = b_{30} + b_{32}y_{t1} + \varepsilon_{t3}. \end{cases}$$

В табл.11 приведены данные наблюдений эндогенных y_{ti} и экзогенных x_{ti} переменных модели:

Таблица 11

t	y_{t1}	y_{t2}	y_{t3}	x_{t1}	x_{t2}	x_{t3}
1	19	23	13	5	5	4
2	15	31	15	7	8	3
3	22	13	17	4	8	2
4	6	32	9	6	2	5
5	8	66	16	12	2	9

Необходимо, опираясь на данные наблюдений, оценить параметры структурной системы.

Определим, являются ли уравнения системы идентифицируемыми. Для этого применим необходимое условие идентификации или счетное правило:

D+1=H – уравнение идентифицируемо;

D+1<H – уравнение неидентифицируемо;

D+1>H – уравнение сверхидентифицируемо.

Здесь - H – число эндогенных переменных в уравнении; D – число отсутствующих экзогенных и лаговых переменных в уравнении. В приведенной системе лаговые или запаздывающие переменные отсутствуют.

В первом уравнении эндогенных переменных две: y_{t1}, y_{t2}, следовательно, H=2. Всего у нас три экзогенных переменных, из них в первом уравнении отсутствует одна x_{t3}. Поэтому D=1. Для первого уравнения выполняется D+1=H=2. Первое уравнение идентифицируемо. Аналогично определяем, что для второго уравнения выполняется D+1=H=3. Для третьего уравнения находим, что D+1=4>H=2. Третье уравнение сверхидентифицируемо.

Прямая оценка параметров структурной системы уравнений недопустима. Из-за корреляции регрессоров с остатками получаются смещенные оценки. Для оценки параметров используют косвенный МНК (в случае идентификации) или двухшаговый МНК (в случае идентификации и сверхидентификации). Оба метода требуют оценки параметров приведенной формы модели:

$$\begin{cases} \hat{y}_{t1} = \hat{c}_{10} + \hat{c}_{11}x_{t1} + \hat{c}_{12}x_{t2} + \hat{c}_{13}x_{t3} \\ \hat{y}_{t2} = \hat{c}_{20} + \hat{c}_{21}x_{t1} + \hat{c}_{22}x_{t2} + \hat{c}_{23}x_{t3} \\ \hat{y}_{t1} = \hat{c}_{30} + \hat{c}_{31}x_{t1} + \hat{c}_{32}x_{t2} + \hat{c}_{33}x_{t3}. \end{cases}$$

Опираясь на данные наблюдений, выполняем регрессию y_{t1} на x_{t1}, x_{t2}, x_{t3}. Получаем уравнение прогноза

$$\hat{y}_{t1} = \underset{(2{,}726)}{-14{,}982} - \underset{(2{,}726)}{5{,}421}\, x_{t1} + \underset{(0{,}367)}{5{,}228}\, x_{t2} + \underset{(0{,}809)}{8{,}632}\, x_{t3} \qquad R^2 = 0{,}998$$

В скобках приводятся стандартные ошибки оценок параметров.

Затем выполняем регрессию y_{t2} на x_{t1}, x_{t2}, x_{t3}

$$\hat{y}_{t2} = \underset{(2,726)}{1,129} + \underset{(0,469)}{5,912}\, x_{t1} - \underset{(0,367)}{1,327}\, x_{t2} - \underset{(0,809)}{0,368}\, x_{t3} \qquad R^2 = 0,9997$$

и y_{t3} на x_{t1}, x_{t2}, x_{t3}

$$\hat{y}_{t3} = \underset{(7,496)}{-\,7,298} - \underset{(1,291)}{1,342}\, x_{t1} + \underset{(1,008)}{2,623}\, x_{t2} + \underset{(2,223)}{3,763}\, x_{t3} \qquad R^2 = 0,920$$

Приведенная система эконометрических уравнений имеет вид

$$\begin{cases} \hat{y}_{t1} = -14,982 - 5,421x_{t1} + 5,228x_{t2} + 8,632x_{t3} \\ \hat{y}_{t2} = 1,129 + 5,912x_{t1} - 1,327x_{t2} - 0,368x_{t3} \\ \hat{y}_{t3} = -7,298 - 1,342x_{t1} + 2,623x_{t2} + 3,763x_{t3}. \end{cases}$$

Параметры первого уравнения структурной системы можно идентифицировать методом, который называется косвенный МНК. Для этого второе уравнение приведенной системы решается относительно x_{t3}:

$$x_{t3} = (\hat{y}_{t2} - 1,129 - 5,912x_{t1} + 1,327x_{t2})/(-0,368) = 3,068 - 2,717\hat{y}_{t2} + 16,06$$

Затем найденное значение подставляется в первое уравнение:

$$\hat{y}_{t1} = -14,982 - 5,421x_{t1} + 5,228x_{t2} + 8,632 \cdot (3,068 - 2,717\hat{y}_{t2} + 16,065x_{t1} -$$

$$= 11,501 - 23,453\hat{y}_{t2} + 133,252x_{t1} - 25,899x_{t2}.$$

Для идентификации второго уравнения структурной системы решаем первое и третье уравнения приведенной системы относительно x_{t1} и x_{t2} и подставляем найденные значения во второе уравнение:

$$\begin{cases} -5{,}421x_{t1} + 5{,}228x_{t2} = \hat{y}_{t1} + 14{,}982 - 8{,}632x_{t3} \\ -1{,}342x_{t1} + 2{,}623x_{t2} = \hat{y}_{t3} + 7{,}298 - 3{,}763x_{t3} \end{cases}$$

$$x_{t1} = \frac{\Delta_1}{\Delta} = \frac{\begin{vmatrix} \hat{y}_{t1} + 14{,}982 - 8{,}632x_{t3} & 5{,}228 \\ \hat{y}_{t3} + 7{,}298 - 3{,}763x_{t3} & 2{,}623 \end{vmatrix}}{\begin{vmatrix} -5{,}421 & 5{,}228 \\ -1{,}342 & 2{,}623 \end{vmatrix}} = -0{,}159 - 0{,}364\hat{y}_{t1} + 0{,}726\hat{y}_{t3} +$$

$$x_{t2} = \frac{\Delta_2}{\Delta} = \frac{\begin{vmatrix} -5{,}421 & \hat{y}_{t1} + 14{,}982 - 8{,}632x_{t3} \\ -1{,}342 & \hat{y}_{t3} + 7{,}298 - 3{,}763x_{t3} \end{vmatrix}}{\begin{vmatrix} -5{,}421 & 5{,}228 \\ -1{,}342 & 2{,}623 \end{vmatrix}} = 2{,}701 - 0{,}186\hat{y}_{t1} + 0{,}753\hat{y}_{t3} -$$

$$\begin{aligned} \hat{y}_{t2} &= 1{,}129 + 5{,}912 \cdot (-0{,}159 - 0{,}364\hat{y}_{t1} + 0{,}726\hat{y}_{t3} + 0{,}412x_{t3}) - \\ &- 1{,}327 \cdot (2{,}701 - 0{,}186\hat{y}_{t1} + 0{,}753\hat{y}_{t3} - 1{,}224x_{t3}) - 0{,}368x_{t3} = \\ &= -3{,}395 - 1{,}905\hat{y}_{t1} + 3{,}293\hat{y}_{t3} + 3{,}692x_{t3}. \end{aligned}$$

Третье уравнение структурной системы сверхидентифицируемо. Для идентификации его параметров необходимо применить двухшаговый МНК. На первом шаге определяем значения инструментальной переменной

$$\hat{y}_{t1} = -14{,}982 - 5{,}421x_{t1} + 5{,}228x_{t2} + 8{,}632x_{t3}.$$

Значения $\hat{y}_{t1}$ рассчитываются в табл.12.

Таблица 12

t	y_{t3}	x_{t1}	x_{t2}	x_{t3}	$\hat{y}_{t1}$
1	13	5	5	4	18,579
2	15	7	8	3	14,789

3	17	4	8	2	22,421
4	9	6	2	5	6,105
5	16	12	2	9	8,105

На втором шаге осуществляем регрессию y_{t3} на $\hat{y}_{t1}$. Получаем уравнение

$$\hat{y}_{t3} = \underset{(3,348)}{10,370} + \underset{(0,219)}{0,259}\,\hat{y}_{t1} \qquad R^2 = 0,319$$

Идентификация параметров произведена. Структурная система эконометрических уравнений имеет вид:

$$\begin{cases} \hat{y}_{t1} = 11,501 - 23,453\hat{y}_{t2} + 133,252x_{t1} - 25,899x_{t2} \\ \hat{y}_{t2} = -3,395 - 1,905\hat{y}_{t1} + 3,293\hat{y}_{t3} + 3,692x_{t3} \\ \hat{y}_{t3} = 10,370 + 0,259\hat{y}_{t1} + \varepsilon_{t3}. \end{cases}$$

3.5. Временные ряды. Авторегрессия

В табл.13 приводятся наблюдения продаж некоторого товара. Переменная t - номер наблюдения. Переменная y_t – выручка за один день.

Таблица 13

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
y_t	19	15	20	21	21	19	16	20	20	22	19	16	22	22	23
t	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
y	2	1	2	2	2	2	1	2	2	2	2	1	2	2	2

t	1	9	3	1	3	2	9	1	1	3	1	9	2	1	4

1. Данные таблицы представляют временной ряд. Оценим характер временного ряда визуально. Для этого построим его диаграмму, представленную на рис.3.

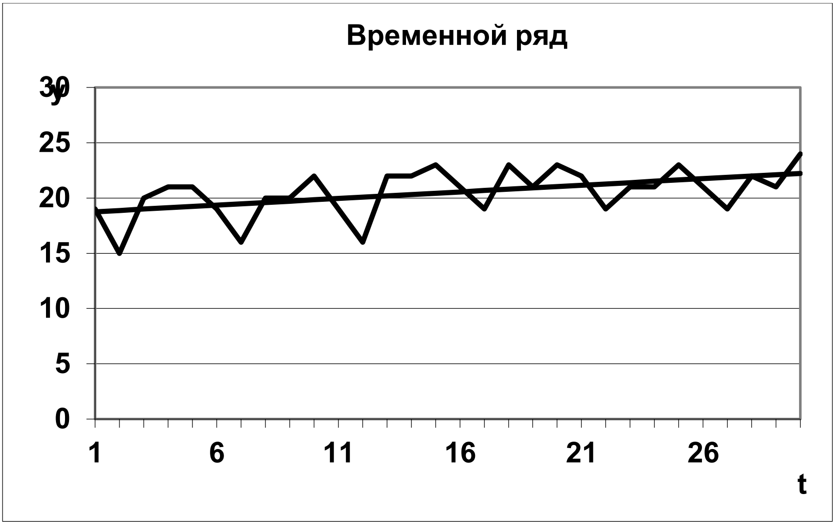

Рис.3. Временной ряд и тренд (прямая линия) продаж товара.

По графику временного ряда можно увидеть, что значения временного ряда возрастают с увеличением номера наблюдения. Это означает, что у временного ряда может быть тренд. Определим параметры тренда. Для этого построим линейную регрессионную модель с регрессором t:

$$y_t = a + b \cdot t + \varepsilon_t .$$

Определим оценки параметров: $\widehat{a} = 18{,}641\ (t = 25{,}932)$; $\widehat{b} = 0{,}120\ (t = 2{,}961)$ (в скобках представлено отношение Стьюдента, критическое значение t_a=2,048, число степеней свободы 30-2=28). Поскольку 2,961>2,048, оценка параметра b является значимой и тренд у временного ряда существует.

Вычислим прогнозные значения

$$\widehat{y}_t = \widehat{a} + \widehat{b} \cdot t = 18{,}641 + 0{,}120 \cdot t$$

и отклонения от прогноза

$$e_t = y_t - \widehat{y}_t.$$

Результаты расчетов представлены в табл.14. Линия тренда показана на рис.3.

Сумма квадратов отклонений остатков построенной модели ESS=103,2, общая сумма квадратов TSS=135,5. Коэффициент детерминации R^2=0,239, отношение Фишера F=8,771 (критическое значение F_a=4,196). Модель тренда является значимой по критерию Фишера.

2. Исследуем возможности увеличения значимости данной модели. Для этого определим выборочную автокорреляционную функцию остатков построенной модели. Ее значения рассчитываются по формуле

$$r_i \approx \frac{\sum_{t=i+1}^{n} e_t e_{t-i}}{\sum_{t=1}^{n} e_t^2}$$

и представлены в табл.15. Гистограмма автокорреляционной функции показана на рис.4.

Таблица 15

Автокорреляционная функция							
i	0	1	2	3	4	5	6
r_i	1	-0,015	-0,273	-0,336	-0,04	0,677	-0,094

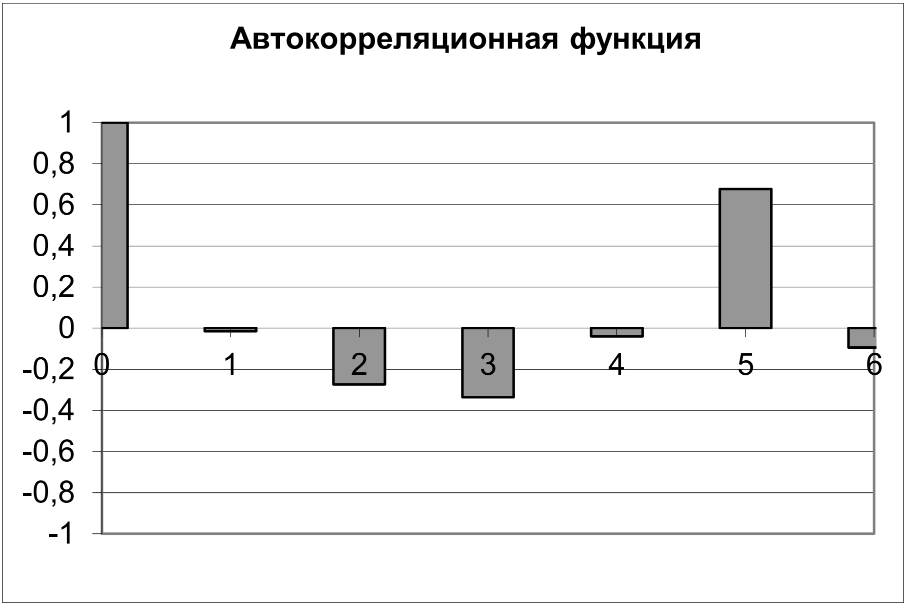

Рис.4. Гистограмма автокорреляционной функции.

При i=5 автокорреляция (0,677) значительно превосходит по модулю соседние значения, и указывает на корреляцию остатков через 5 наблюдений. Для продаж, которые осуществляются по рабочим дням (5 дней в неделю), характерна так называемая недельная сезонная компонента. Если фиксируются поквартальные данные, то может наблюдаться корреляция остатков при i=4 (квартальная сезонная компонента). Если рассматриваются месячные данные, то может наблюдаться значимая корреляция при i=12 и т.д.

В связи с тем, что отклонения продаж товара от тренда, коррелируют через 5 дней, предлагается следующая модель временного ряда:

$$y_t = a + b \cdot t + \varepsilon_t$$

$$\varepsilon_t = \rho \cdot \varepsilon_{t-5} + u_t,$$

где u_t – случайные отклонения, удовлетворяющие требованиям классической ЛРМ.

Вычислим

$$y_t^* = y_t - \rho y_{t-5} = a + bt + \varepsilon_t - \rho(a + b(t-5) + \varepsilon_{t-5}) =$$

$$= a(1-\rho) + b(t - \rho(t-5)) + \varepsilon_t - \rho\varepsilon_{t-5} = a^* + bt^* + u_t$$

Получена классическая ЛРМ, параметры которой могут оцениваться с помощью МНК при этом оценки являются состоятельными, несмещенными и эффективными. Оценка параметра $\hat{\rho} = r_5 = 0{,}677$. Вычислим $y_t^* = y_t - \hat{\rho} y_{t-5}$ и $t^* = t - \hat{\rho}(t-5)$ и включим найденные значения в таблицу. МНК оценки параметров $\hat{a}^* = 6{,}332$; $\hat{a} = \dfrac{\hat{a}^*}{1-\hat{\rho}} = \dfrac{6{,}332}{1-0{,}677} = 19{,}603$; $\hat{b} = 0{,}073$.

Получена ЛРМ с учетом сезонных колебаний

$$\hat{y}_t = \hat{a} + \hat{b}t = 19{,}603 + 0{,}073 \cdot t.$$

Прогнозирование в этой модели осуществляется с учетом последних наблюдений временного ряда:

$$\hat{y}_t^{пр} = \hat{y}_t^* + \hat{\rho} y_{t-5} = \hat{y}_t + \hat{\rho}(y_{t-5} - \hat{y}_{t-5}).$$

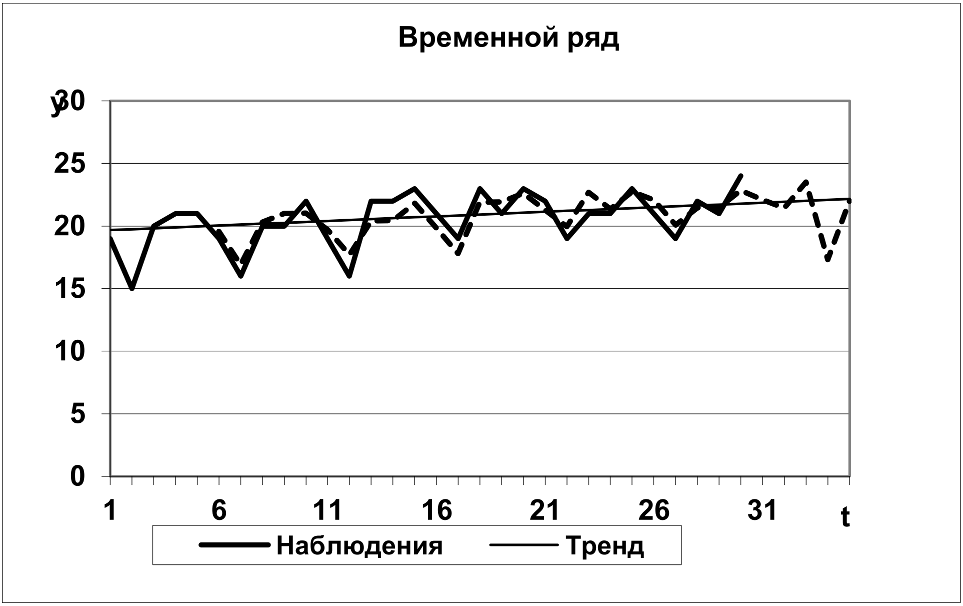

Рис.5. Временной ряд, линейный тренд и прогноз с учетом корреляции остатков.

Сумма квадратов остатков построенной модели ESS=26,5, общая сумма квадратов TSS=135,5. Коэффициент детерминации R^2=0,804, отношение Фишера F=94,6 (критическое значение F_a=4,279, число степеней свободы 30-2-5=23). Значимость модели заметно улучшилась. Прогноз продаж товара на следующую рабочую неделю представлен в таблице и на рис.5.