Федеральное государственное бюджетное образовательное учреждение высшего образования «Кемеровский государственный университет» Новокузнецкий институт (филиал)

Факультет информатики, математики и экономики

Кафедра экономики и управления

М.А. Кречетова

ЭКОНОМЕТРИКА

Методические указания по выполнению контрольной работы

для обучающихся по специальности

38.05.01 Экономическая безопасность

Форма обучения – заочная

Новокузнецк

2020

УДК 330.4 (075.8)

ББК 65в6я73

Кречетова М.А.

Эконометрика: метод. указ. по выполнению контрольной работы по специальности 38.05.01 Экономическая безопасность, заочной формы обучения/ М.А. Кречетова. - Новокузнецк ин-т (фил.) Кемеров. гос. ун-та. – Новокузнецк: НФИ КемГУ, 2020. – 50 с. - Текст: непосредственный.

В методических указаниях для студентов представлены рекомендации по выполнению контрольной работы по дисциплине «Эконометрика»: требования к выполнению контрольной работы; перечень заданий, критерии оценивания контрольной работы, список литературы.

Методические указания предназначены для организации самостоятельной работы студентов формы обучения по специальности 38.05.01 Экономическая безопасность.

Рекомендовано
на заседании кафедры экономики
и управления
4 декабря 2020 г.
Заведующий кафедрой
______________ Ю. Н. Соина-Кутищева

Утверждено
методической комиссией
факультета математики,
информатики и экономики
«___»_________________
Председатель методкомиссии
____________Г.Н. Бойченко

Текст представлен в авторской редакции.

ВВЕДЕНИЕ

Цель изучения дисциплины «Эконометрика» - подготовка специалистов, владеющих современной методологией эконометрических моделей для анализа и моделирования показателей рыночной экономики.

Учебная дисциплина «Эконометрика» является базовой для специальности 38.05.03 Экономическая безопасность, формирующей навыки владения эконометрической методологией экономистов любого профиля, в том числе и специалистов по безопасности.

В современных условиях навыки построения и анализа эконометрических моделей и прогнозов являются непременным условием принятия обоснованных управленческих решений на будущий период развития. Знание эконометрической методологии и умение ее применять позволяет эффективно оценивать текущую экономическую и социальную ситуацию, спрогнозировать ее развитие в будущем и определить наиболее оптимальные пути решения различных задач в сфере экономики и безопасности.

В результате освоения дисциплины у обучающегося должна быть сформирована **компетенции** основной профессиональной образовательной программы специалитета;

ОПК-1	способностью применять математический инструментарий для решения экономических задач

Требуемая подготовка: знание основ экономических знаний, макроэкономики, высшей математики и статистики.

Цель настоящих методических указаний – дать студентам возможность углубить и закрепить полученные на лекциях и при самостоятельном изучении дисциплины знания, а также научить их применять эти знания на практике при решении различных задач в сфере безопасности. Предлагаемые методические указания адресованы студентам заочного и очно-заочного отделения ВУЗов, обучающихся по специальности 38.05.03 Экономическая безопасность.

Методические указания содержат задания по основным темам курса, методические указания по их решению и примеры решения некоторых задач, требования к выполнению контрольной работы.

1. Основные требования к выполнению контрольной работы

1.1 Выбор варианта работы

Контрольная работа представлена в десяти вариантах, номер варианта определяется по **последней цифре** в зачетке студента.

Приступая к выполнению контрольной работы, необходимо ознакомиться с соответствующими разделами программы курса и методическими указаниями, изучить литературу. Особое внимание нужно обратить на методы построения эконометрических моделей, методы оценки их качества по имеющимся статистическим данным, характеризующим изучаемые экономические явления и процессы.

1.2 Требования по выполнению расчетных заданий контрольной работы

Общие требования

1. Контрольная работа должна быть выполнена и представлена в срок, установленный преподавателем.
2. В начале работы должен быть указан номер варианта работы. Задачи нужно решать в том порядке, в каком они даны в задании.
3. Решение задач следует сопровождать необходимыми формулами, развернутыми расчетами и пояснениями. Если имеется несколько методов расчета того или иного показателя, надо применять наиболее простой из них, указав при этом другие способы решения.
4. Решение задач следует по возможности оформлять в виде таблиц и графиков.
5. В конце решения каждой задачи необходимо четко сформулировать выводы. Без вывода задача считается нерешенной.
6. Все расчеты относительных показателей нужно производить с принятой в статистике точностью до 0,001, а проценты - до 0,1.
7. Страницы работы должны быть пронумерованы, и иметь достаточно широкие поля для замечаний рецензента и исправлений (дополнений), вносимых студентом после рецензирования.
8. В конце работы следует привести список использованной литературы. Работа должна быть подписана студентом с указанием даты ее выполнения.
9. Выполненная работа сдается на проверку преподавателю. Удовлетворительно выполненная работа оценивается на «зачтено». Студенты, представившие на проверку неудовлетворительные работы, выполняют работу заново **с** учетом замечаний преподавателя. На зачет допускаются студенты с зачтенными работами.

Порядок оценивания

10. Контрольная работа оценивается на «зачтено» и «не зачтено». Минимальное количество баллов за контрольную работу 32, максимальное 48. Зачтено ставится студенту, если он набрал 32 балла и выше. Правильное выполнение каждой задачи с подробными выводами оценивается максимальным количеством баллов, решение с недочетами меньшим количеством баллов. За соблюдение требований к оформлению добавляется 2 балла.

Критерии оценивания	Баллов всего	
	Мин.	Макс.
Задание 1 - Парная регрессия	6	10
Задание 2 – Нелинейная регрессия	8	12
Задание 3 - Построение тренда и адаптивные методы	8	12
Задание 4 - Статистический анализ динамики и сезонность	10	14
Выполнение требований к оформлению	0	2
Итого	**32**	**48**

1.3 Методические указания по содержанию и оформлению контрольной работы

В ходе выполнения контрольной работы студент выполняет самостоятельную работу, посвященную изучению и закреплению конкретного направления теории и практики статистического анализа данных, что способствует решению ряда важнейших задач обучения, а именно систематизации, закреплению и расширению теоретических знаний и практических навыков статистического анализа, применение этих знаний при решении конкретных расчетных заданий, рассматриваемых в контрольной работе.

Структура контрольной работы

Структура контрольной работы следующая:

1 Титульный лист

2 Содержание.

3 Решение расчетных заданий.

4 Список использованных источников.

Титульный лист

На титульном листе должны быть указаны имя, фамилия, курс, группа, факультет, направление, профиль и номер варианта (Приложение 1).

Содержание работы

Содержание включает наименование всех разделов, подразделов и пунктов с указанием номера страниц, с которых они начинаются.

Решение расчетных заданий

Основная часть контрольной работы - решение расчетных заданий, подкрепленное формулами, таблицами; обоснование самостоятельных выводов и предложений.

За качество, правильность представленных расчетов, обоснованные решения и рекомендации, а также за своевременное и в соответствующем объеме выполнение контрольной работы несет ответственность студент.

Список использованных источников

Список литературы, составляемый студентом, должен включать только источники, непосредственно использованные в контрольной работе, т.е. те, которые цитировались, на которые делались ссылки или если они послужили основой для формирования точки зрения студента.

Правила оформления работы

Работа представляется к защите в виде текста, к которому предъявляется ряд требований по оформлению.

Оформление текста контрольной работы выполняется в соответствии с требованиями, изложенными в методических указаниях - Правила оформления учебных работ студентов: учебно-методическое пособие / Новокузнец. ин-т (фил.) Кемеров. гос. ун-та ; сост.: И. А. Жибинова [и др.]; под ред. И. А. Жибиновой. – Новокузнецк: НФИ КемГУ, 2018. – 104 с. Литература оформляется в соответствии с ГОСТ.

2 Задания на контрольную работу
Задание 1. Парная регрессия

1. Построить график – ***поле корреляции*** и сделать вывод о виде модели.
2. Построить линейное уравнение парной регрессии y от x.
3. Рассчитать линейный коэффициент парной корреляции, коэффициент детерминации и среднюю ошибку аппроксимации. Сделать выводы.
4. Оценить статистическую значимость параметров регрессии и корреляции с помощью F-критерия Фишера и t-критерия Стьюдента. Сделать выводы.
5. На одном графике построить исходные данные и теоретическую прямую.
6. Выполнить прогноз y* при прогнозном значении x*, заданном в задании
7. Оценить точность прогноза, рассчитав ошибку прогноза и его доверительный интервал.

ЗАДАНИЯ по вариантам

1. Туристическая компания предлагает места в гостиницах приморского курорта. Менеджера компании интересует, насколько возрастает привлекательность гостиницы в зависимости от ее удаленности от пляжа. С этой целью по 10 гостиницам города была выяснена среднегодовая наполняемость номеров и расстояние от пляжа:

Расстояние, км X	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,7	0,8	0,9
Наполняемость, % Y	95	96	90	89	90	83	85	80	78	75

К пункту 6. Значение x^* =0,75.

2. Компанию по прокату автомобилей интересует зависимость между пробегом автомобилей и стоимостью ежемесячного обслуживания. Для выяснения характера этой связи было отобрано 10 автомобилей. Результаты исследования представлены в таблице:

Пробег, тыс. км, X	6	7	8	9	10	11	12	13	14	15
Стоимость обслуживания, долл., Y	13	16	15	20	19	21	26	23	30	32

К пункту 6. Значение x^* =18.

3. Торговцу нужно выяснить, как изменяется количество пучков салата, продаваемых ежедневно в розницу. Имеются следующие сведения о количестве и цене:

Цена, у.е. за единицу, X	30	31	25	26	22	24	16	12	10	8
Количество, тыс./день, Y	28	29	34	35	37	37	41	46	48	52

К пункту 6. Значение x^* =20.

4. Компания, занимающаяся продажей радиоаппаратуры, установила на видеомагнитофон определенной модели цену, дифференцированную по регионам. Следующие данные показывают цены на видеомагнитофон в 10 различных регионах и соответствующее им число продаж:

Цена, у.е., X	5,5	6,0	6,5	6,0	5,0	5,6	4,5	5,0	4,4	4,0
Число продаж, шт., Y	420	380	350	400	440	380	450	420	460	500

К пункту 6. Значение x^* =5,7.

5. Некоторая компания недавно провела рекламную кампанию в магазинах с демонстрацией антисептических качеств своего нового моющего средства. Через 10 недель компания решила проанализировать эффективность этого вида рекламы, сопоставив еженедельные объемы продаж с расходами на рекламу (тыс. руб.):

Расходы на рекламу, у.е., X	5	8	6	5	3	9	12	4	3	10
Объем продаж, у.е., Y	72	76	78	70	68	80	82	65	62	90

К пункту 6. Значение x^* =7

6. По 10 однородным предприятиям имеются данные о количестве рабочих с профессиональной подготовкой и количестве бракованной продукции:

Количество рабочих с профессиональной подготовкой, %, X	10	12	14	17	24	28	30	35	40	50
Количество бракованной продукции, %, Y	18	17	14	12	10	10	8	9	6	6

К пункту 6. Значение x^* =27.

7. При исследовании годового дохода и сбережений населения в случайном порядке отобрано 10 человек. Получены следующие данные:

Доход, тыс. у.е., X	15	6	9	3	20	11	14	10	12	14
Сбережения, у.е., Y	2000	200	500	500	2500	1800	1500	1500	1600	1800

К пункту 6. Значение x^* =17.

8. Проведен опрос случайно выбранных 10 студентов, проживающих в общежитии университета, для выявления зависимости между средним баллом по результатам предыдущей сессии и числом часов в неделю, затраченных студентом на самостоятельную подготовку:

Число часов. X	25	22	19	15	15	30	20	30	10	17
Средний балл, Y	4,6	4,3	3,8	3,8	4,2	4,3	3,8	4,0	3,1	3,9

К пункту 6. Значение x^* =27.

9. Имеются данные по 10 предприятиям о производительности труда и коэффициенте механизации работ:

Коэффициент механизации работ, %, X	36	40	41	47	56	54	60	55	61	67
Производительность труда, шт., Y	28	30	31	33	34	37	38	40	41	43

К пункту 6. Значение x^* =65.

10. Представлены данные, характеризующие зависимость между % лишнего веса и количеством больных диабетом по району:

Количество больных с сахарным диабетом, Y	30	33	45	48	53	25	29	34	65	25	28	35
Лишний вес, % , X	65	60	70	77	80	50	45	66	90	40	48	59

К пункту 6. Значение x^* = 30.

Задание 2. «Нелинейная регрессия»

По предприятиям промышленности региона получена информация, характеризующая зависимость ***объема выпуска продукции (Y, млн. руб.)*** от ***объема капиталовложений (X, млн. руб.).*** Требуется:

1. Для характеристики зависимости *Y* от *X* построить следующие модели (с помощью графического анализа):

—линейную, степенную, параболу, логарифмическую, экспоненту.

2. Оценить каждую модель, определив:

— коэффициент детерминации,

— *F*-критерий Фишера

– ошибку аппроксимации,

– коэффициент эластичности.

Дать интерпретацию рассчитанных характеристик для каждой модели.

3. Составить сводную таблицу вычислений, обосновать выбор лучшей модели.

4. Рассчитать прогнозные значения результативного признака Y, если прогнозное значение фактора X увеличится на 20 % относительно максимального значения.

Задания по вариантам

Вариант 1

X	66	58	73	82	81	84	55	67	81	59
Y	133	107	145	162	163	170	104	132	159	116

Вариант 2

X	72	52	73	74	76	79	54	68	73	64
Y	121	84	119	117	129	128	102	111	112	98

Вариант 3

X	38	28	27	37	46	27	41	39	28	44
Y	69	52	46	63	73	48	67	62	47	67

Вариант 4

X	36	28	43	52	51	54	25	37	51	29
Y	104	77	117	137	143	144	82	101	132	77

Вариант 5

X	31	23	38	47	46	49	20	32	46	24
Y	38	26	40	45	51	49	34	35	42	24

Вариант 6

X	33	17	23	17	36	25	39	20	13	12
Y	43	27	32	29	45	35	47	32	22	24

Вариант 7

X	36	28	43	52	51	54	25	37	51	29
Y	85	60	99	117	118	125	56	86	115	68

Вариант 8

X	17	22	10	7	12	21	14	7	20	3
Y	26	27	22	19	21	26	20	15	30	13

Вариант 9

X	12	4	18	27	26	29	1	13	26	5
Y	21	10	26	33	34	37	9	21	32	14

Вариант 10

X	26	18	33	42	41	44	15	27	41	19
Y	43	28	51	62	63	67	26	43	61	33

Задание 3. Построение трендов. Адаптивные методы.

Дана выборка динамики производства продукции (тыс. ед.) за 12 лет по предприятию (таблица 1).

1) Провести сглаживание временного ряда с помощью трехчленной скользящей средней. Сделать выводы о типе модели.
2) Построить модель тенденции временного ряда (с помощью подбора в Excel). Выбрать лучшую по скорректированному коэффициенту детерминации, рассчитать значения по модели.
3) Оценить качество модели (по коэффициенту детерминации, критерию поворотных точек, ошибке аппроксимации, критерию Дарбина-Уотсона, критерию RS). Сделать выводы о качестве и возможности прогнозирования по модели.
4) Построить прогноз на 2020 г. по лучшей модели (t* = 14).

5) Рассчитать адаптивную модель Брауна для параметров α = 0,5; 0,2; 0,8 для каждой модели найти стандартную ошибку и выбрать лучшую. Построить прогноз по модели на 2020 г.

Таблица 1. Данные для анализа по вариантам

год	Вариант									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
2007	20	25	200	115	105	38	240	102	15	200
2008	25	29	183	105	130	41	232	97	18	143
2009	27	27	153	95	150	42	214	89	15	140
2010	30	36	140	92	250	49	200	77	30	170
2011	29	48	107	84	242	64	230	81	35	250
2012	30	73	87	72	254	53	280	87	26	280
2013	35	68	68	57	275	47	300	94	34	340
2014	33	77	83	49	272	52	310	92	37	354
2015	40	78	85	46	262	51	330	100	44	370
2016	43	84	79	40	259	70	290	104	59	320
2017	42	87	80	35	330	92	237	87	64	250
2018	48	80	73	32	362	100	252	80	68	210

Задание 4. Прогнозирование сезонности.

Дана выборка объемов продаж разных товаров по одному из магазинов за 12 месяцев (таблица 2).

1) Найти коэффициенты автокорреляции со смещением на 1,2,3,4 и их значимость с вероятностью 0,95. Сделать выводы.
2) Построить коррелограмму.
3) Провести сглаживание ряда трехчленными скользящими средними.
4) Определить скорректированную сезонную компоненту. И исключить ее из ряда.
5) Построить аддитивную (или мультипликативную) модель временного ряда.
6) Оценить качество модели по коэффициенту детерминации, ошибке аппроксимации и RS. Сделать выводы.
7) Построить прогноз по модели на февраль следующего года.

Таблица 2 Данные для анализа по вариантам

месяц	***№ варианта***									
	1	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***
1	52	79	74,4	107	84,1	112	32,8	46,7	13,3	35
2	52,1	78,2	73,2	108	82,6	111	32	46,1	12,5	33
3	57,3	78,6	74,3	111	83,8	112	32,8	45,7	12,7	34
4	56,5	82	79,9	110	87,5	117	34,7	49,7	17,2	38
5	56,1	81	78,7	109	87,3	116	34,1	47,4	15,9	36
6	56,2	82,3	79,7	112	88,1	116	34,2	47,8	16,1	38

7	61,3	87,1	84,1	117	93	122	37,5	52	20,5	42
8	60,9	86,3	84,3	116	92,3	121	35,8	50,1	19,2	40
9	60,5	85,5	85,4	117	93,6	121	35,7	49,8	19,9	41
10	65,4	92,4	89,3	122	99,4	126	40,1	54,6	23,9	45
11	65,6	90,6	89,6	121	97,2	125	38,8	51,9	22,8	44
12	65	90,7	91	122	98	126	39	52,3	23,5	45

3. Методические указания по выполнению контрольной работы

3.1 Типовой пример «Парная регрессия»

По данным проведенного опроса восьми групп семей известны данные связи расходов населения на продукты питания с уровнем доходов семьи.

Таблица 3. Данные по доходам и расходам семей

Расходы на продукты питания, тыс. руб., Y	0,9	1,2	1,8	2,2	2,6	2,9	3,3	3,8
Доходы семьи, тыс. руб.. X	1,2	3,1	5,3	7,4	9,6	11,8	14,5	18,7

1) Построить график – ***поле корреляции*** и сделать вывод о виде модели.
2) Рассчитать параметры парной регрессии y от x.
3) Рассчитать линейный коэффициент парной корреляции, коэффициент детерминации и среднюю ошибку аппроксимации. Сделать выводы.
4) Оценить статистическую значимость параметров регрессии и корреляции, а также всего уравнения с помощью F-критерия Фишера и t-критерия Стьюдента. Сделать выводы.
5) На одном графике построить исходные данные и теоретическую прямую.
6) Выполнить прогноз y^* при прогнозном значении x^*, равном 110% от максимального значения. Оценить точность прогноза, рассчитав ошибку прогноза и его доверительный интервал.

РЕШЕНИЕ: Предположим, что связь между доходами семьи и расходами на продукты питания линейная. Для подтверждения нашего предположения построим поле корреляции.

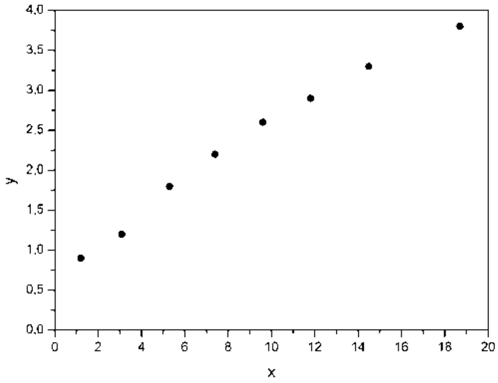

Рис. 3.1 - Поле корреляции

По графику видно, что точки выстраиваются в некоторую прямую линию.

Для удобства дальнейших вычислений составим таблицу 4.

Таблица 4. Рабочая таблица

	x	y	$x \cdot y$	x^2	y^2	y_x	$y - y_x$	$\left(y - y_x\right)^2$	A_i, %
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	1,2	0,9	1,08	1,44	0,81	1,038	–0,138	0,0190	15,33
2	3,1	1,2	3,72	9,61	1,44	1,357	–0,157	0,0246	13,08
3	5,3	1,8	9,54	28,09	3,24	1,726	0,074	0,0055	4,11
4	7,4	2,2	16,28	54,76	4,84	2,079	0,121	0,0146	5,50
5	9,6	2,6	24,96	92,16	6,76	2,449	0,151	0,0228	5,81
6	11,8	2,9	34,22	139,24	8,41	2,818	0,082	0,0067	2,83
7	14,5	3,3	47,85	210,25	10,89	3,272	0,028	0,0008	0,85
8	18,7	3,8	71,06	349,69	14,44	3,978	–0,178	0,0317	4,68
Итого	71,6	18,7	208,71	885,24	50,83	18,717	–0,017	0,1257	52,19
Среднее значение	8,95	2,34	26,09	110,66	6,35	2,34	–	0,0157	6,52
σ	5,53	0,935	–	–	–	–	–	–	–
σ^2	30,56	0,874	–	–	–	–	–	–	–

Рассчитаем параметры линейного уравнения парной регрессии $y_x = a + b \cdot x$. Для этого воспользуемся формулами (1.5):

$$b = \frac{\operatorname{cov}(x,\ y)}{\sigma_x^2} = \frac{\overline{x \cdot y} - \overline{x} \cdot \overline{y}}{\overline{x^2} - \overline{x}^2} = \frac{26,09 - 8,95 \cdot 2,34}{30,56} = 0,168;$$

$$a = \overline{y} - b \cdot \overline{x} = 2,34 - 0,168 \cdot 8,95 = 0,836.$$

Получили уравнение: $y_x = 0{,}836 + 0{,}168 \cdot x$. Т.е. с увеличением дохода семьи на 1000 руб. расходы на питание увеличиваются на 168 руб.

Как было указано выше, уравнение линейной регрессии всегда дополняется показателем тесноты связи – линейным коэффициентом корреляции r_{xy}:

$$r_{xy} = b \cdot \frac{\sigma_x}{\sigma_y} = 0{,}168 \cdot \frac{5{,}53}{0{,}935} = 0{,}994 \cdot$$

Близость коэффициента корреляции к 1 указывает на тесную линейную связь между признаками.

Коэффициент детерминации $r_{xy}^2 = 0{,}987$ показывает, что уравнением регрессии объясняется 98,7% дисперсии результативного признака, а на долю прочих факторов приходится лишь 1,3%.

Оценим качество уравнения регрессии в целом с помощью F-критерия Фишера. Сосчитаем фактическое значение F-критерия:

$$F = \frac{r_{xy}^2}{1 - r_{xy}^2} \cdot (n-2) = \frac{0{,}987}{1 - 0{,}987} \cdot 6 = 455{,}54 \cdot$$

Табличное значение $(k_1 = 1,\ k_2 = n - 2 = 6,\ \alpha = 0{,}05)$ $F_{\text{табл}} = 5{,}99$. Так как $F_{\text{факт}} > F_{\text{табл}}$, то признается статистическая значимость уравнения в целом.

Для оценки статистической значимости коэффициентов регрессии и корреляции рассчитаем t - критерий Стьюдента и доверительные интервалы каждого из показателей. Рассчитаем случайные ошибки параметров линейной регрессии и коэффициента корреляции:

$$\left(S_{\text{ост}}^2 = \frac{\sum (y - y_x)^2}{n-2} = \frac{0{,}1257}{8-2} = 0{,}021 \right)$$

$$m_b = \frac{S_{\text{ост}}}{\sigma_x \cdot \sqrt{n}} = \frac{\sqrt{0{,}021}}{5{,}53 \cdot \sqrt{8}} = 0{,}0093, \qquad m_a = S_{\text{ост}} \cdot \frac{\sqrt{\sum x^2}}{\sigma_x \cdot n} = \frac{\sqrt{0{,}021 \cdot 885{,}24}}{5{,}53 \cdot 8} = 0{,}0975,$$

$$m_r = \sqrt{\frac{1 - r^2}{n-2}} = \sqrt{\frac{1 - 0{,}987}{6}} = 0{,}0465 \cdot$$

Фактические значения t- статистик: $t_b = \frac{0{,}168}{0{,}0093} = 18{,}065$, $t_a = \frac{0{,}836}{0{,}0975} = 8{,}574$, $t_r = \frac{0{,}994}{0{,}0465} = 21{,}376$. Табличное значение t- критерия Стьюдента при $\alpha = 0{,}05$ и числе степеней свободы $\nu = n - 2 = 6$ есть $t_{\text{табл}} = 2{,}447$. Так как $t_b > t_{\text{табл}}$, $t_a > t_{\text{табл}}$ и $t_r > t_{\text{табл}}$, то признаем статистическую значимость параметров регрессии и показателя тесноты связи.

Средняя ошибка аппроксимации находим с помощью столбца 10 таблицы 4:

$$A_i = \left| \frac{y_i - y_{x_i}}{y_i} \right| \cdot 100\%$$

$\bar{A} = 6{,}52\%$ меньше 10%, что говорит о точности построенного уравнения и хорошем качестве уравнения регрессии, т.е. свидетельствует о хорошем подборе модели к исходным данным.

И, наконец, найдем прогнозное значение результативного фактора y_p при значении признака-фактора, составляющем 110% от среднего уровня $x_p = 1{,}1 \cdot \overline{x} = 1{,}1 \cdot 8{,}95 = 9{,}845$, т.е. найдем расходы на питание, если доходы семьи составят 9,85 тыс. руб.

$$y_p = 0{,}836 + 0{,}168 \cdot 9{,}845 = 2{,}490 \text{ (тыс. руб.)}$$

Следовательно, если доходы семьи составят 9,845 тыс. руб., то расходы на питание будут 2,490 тыс. руб.

Найдем доверительный интервал прогноза. Ошибка прогноза

$$m_{y_p} = S_{\text{ост}} \cdot \sqrt{1 + \frac{1}{n} + \frac{\left(x_p - \overline{x}\right)^2}{n \cdot \sigma_x^2}} = \sqrt{0{,}021 \cdot \left(1 + \frac{1}{8} + \frac{(9{,}845 - 8{,}95)^2}{8 \cdot 30{,}56}\right)} = 0{,}154,$$

а доверительный интервал ($y_p - \Delta_{y_p} \le y_p \le y_p + \Delta_{y_p}$)

$$2{,}113 < y_p < 2{,}867.$$

Т.е. прогноз является статистически надежным.

Теперь на одном графике изобразим исходные данные и линию регрессии:

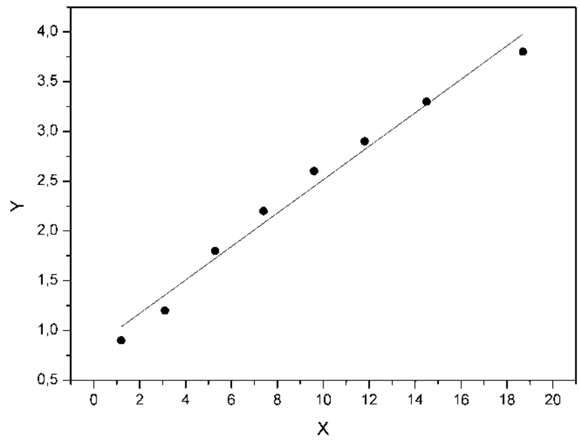

Рис. 2 - Линейная модель

3.2 Нелинейная регрессия

Во многих практических случаях моделирование экономических зависимостей линейными уравнениями дает вполне удовлетворительный результат и может использоваться для анализа и прогнозирования. Однако многие экономические зависимости не являются линейными по своей сути, и поэтому их моделирование линейными уравнениями регрессии не даст положительного результата.

Если между экономическими явлениями существуют нелинейные соотношения, то они выражаются с помощью соответствующих нелинейных функций.

Различают два класса нелинейных регрессий:

1. *Регрессии, нелинейные относительно включенных в анализ объясняющих переменных, но линейные по оцениваемым параметрам, например*

– полиномы различных степеней – $y_x = a + b \cdot x + c \cdot x^2$, $y_x = a + b \cdot x + c \cdot x^2 + d \cdot x^3$;

– равносторонняя гипербола – $y_x = a + b/x$;

– полулогарифмическая функция – $y_x = a + b \cdot \ln x$.

2. *Регрессии, нелинейные по оцениваемым параметрам*, например

– степенная – $y_x = a \cdot x^b$;

– показательная – $y_x = a \cdot b^x$;

– экспоненциальная – $y_x = \mathrm{e}^{a+b\cdot x}$.

Регрессии нелинейные по включенным переменным приводятся к линейному виду простой заменой переменных, а дальнейшая оценка параметров производится с помощью метода наименьших квадратов.

Например, равносторонняя гипербола $y_x = a + b/x$ может быть использована для характеристики связи удельных расходов сырья, материалов, топлива от объема выпускаемой продукции, времени обращения товаров от величины товарооборота, процента прироста заработной платы от уровня безработицы (например, кривая А.В. Филлипса), расходов на непродовольственные товары от доходов или общей суммы расходов (например, кривые Э. Энгеля) и в других случаях. Гипербола приводится к линейному уравнению простой заменой: $z = 1/x$. Система линейных уравнений при применении МНК будет выглядеть следующим образом:

$$\begin{cases} a \cdot n \quad + b \cdot \sum \frac{1}{x} = \sum y; \\ a \cdot \sum \frac{1}{x} + b \cdot \sum \frac{1}{x^2} = \sum \frac{1}{x} \cdot y. \end{cases}$$

Аналогичным образом приводятся к линейному виду зависимости $y_x = a + b \cdot \ln x$, $y_x = a + b \cdot \sqrt{x}$ и другие.

Класс функций, нелинейных по параметрам, в свою очередь, делится на два типа:

- *нелинейные модели внутренне линейные;*
- *нелинейные модели внутренне нелинейные.*

Внутренне линейные модели могут быть приведены к линейному виду.

К внутренне линейным моделям относятся, например,

– степенная функция $y_x = a \cdot x^b$,

– показательная $y_x = a \cdot b^x$,

– экспоненциальная $y_x = \mathrm{e}^{a+b\cdot x}$,

– логистическая $y_x = \dfrac{a}{1 + b \cdot \mathrm{e}^{-c\cdot x}}$,

– обратная $y_x = \frac{1}{a + b \cdot x}$.

1) $y = \alpha \cdot x^{\beta} \cdot \varepsilon$ - внутренне линейна, так как $\ln y = \ln a + b \ln x + \ln \varepsilon$ - линейна по параметрам;

2) $y = e^{a+bx} \cdot \varepsilon$ - внутренне линейна, так как $\ln y = a + bx + \ln \varepsilon$;

3) $y = \frac{1}{a + bx + \varepsilon}$ - внутренне линейна, так как $z = \frac{1}{y} \Rightarrow z = a + bx + \varepsilon$;

4) $y = \frac{a}{1 + be^{-cx+\varepsilon}}$ - логистическая функция – внутренне линейна –

$$\frac{a}{y} = 1 + be^{-cx+\varepsilon}; \quad \frac{a}{y} - 1 = be^{-cx+\varepsilon}; \quad \ln(\frac{a}{y} - 1) = \ln b - cx + \varepsilon.$$

К **внутренне нелинейным** моделям можно, например, отнести следующие модели: $y_x = a + b \cdot x^c$, $y_x = a \cdot \left(1 - \frac{1}{1 - x^b}\right)$, которые к линейному виду не приводятся.

Среди нелинейных моделей наиболее часто используется **степенная функция** $y = a \cdot x^b \cdot \varepsilon$, которая приводится к линейному виду логарифмированием:

$$\ln y = \ln\left(a \cdot x^b \cdot \varepsilon\right);$$

$$\ln y = \ln a + b \cdot \ln x + \ln \varepsilon;$$

$$Y = A + b \cdot X + \mathrm{E},$$

где $Y = \ln y$, $X = \ln x$, $A = \ln a$, $\mathrm{E} = \ln \varepsilon$. Т.е. МНК мы применяем для преобразованных данных:

$$\begin{cases} A \cdot n + b \cdot \sum X = \sum Y, \\ A \cdot \sum X + b \cdot \sum X^2 = \sum X \cdot Y, \end{cases}$$

а затем потенцированием находим искомое уравнение.

Широкое использование степенной функции связано с тем, что параметр b в ней имеет четкое экономическое истолкование – он является коэффициентом эластичности.

Коэффициент эластичности показывает, на сколько процентов измениться в среднем результат, если фактор изменится на 1%.

Формула для расчета коэффициента эластичности имеет вид:

$$Э = f'(x) \cdot \frac{x}{y}.$$

Так как для остальных функций коэффициент эластичности не является постоянной величиной, а зависит от соответствующего значения фактора x, то обычно рассчитывается средний коэффициент эластичности: $\bar{Э} = f'(\bar{x}) \cdot \frac{\bar{x}}{\bar{y}}$.

Приведем формулы для расчета средних коэффициентов эластичности для наиболее часто используемых типов уравнений регрессии:

Таблица 6. Формулы для расчета среднего коэффициента эластичности

Вид функции, y	Первая производная, y'	Средний коэффициент эластичности, $\bar{Э}$
$y = a + b \cdot x + \varepsilon$	b	$\frac{b \cdot \bar{x}}{a + b \cdot \bar{x}}$
$y = a + b \cdot x + c \cdot x^2 + \varepsilon$	$b + 2c \cdot x$	$\frac{(b + 2c \cdot \bar{x}) \cdot \bar{x}}{a + b \cdot \bar{x} + c \cdot \bar{x}^2}$
$y = a + \frac{b}{x} + \varepsilon$	$-\frac{b}{x^2}$	$-\frac{b}{a \cdot \bar{x} + b}$
$y = a \cdot x^b \cdot \varepsilon$	$a \cdot b \cdot x^{b-1}$	b
$y = a \cdot b^x \cdot \varepsilon$	$a \cdot \ln b \cdot b^x$	$\bar{x} \cdot \ln b$
$y = a + b \cdot \ln x + \varepsilon$	$\frac{b}{x}$	$\frac{b}{a + b \cdot \ln \bar{x}}$
$y = \frac{1}{a + b \cdot x + \varepsilon}$	$-\frac{b}{(a + b \cdot x)^2}$	$-\frac{b \cdot \bar{x}}{a + b \cdot \bar{x}}$

Возможны случаи, когда расчет коэффициента эластичности не имеет смысла. Это происходит тогда, когда для рассматриваемых признаков бессмысленно определение изменения в процентах.

Уравнение нелинейной регрессии, так же, как и в случае линейной зависимости, дополняется показателем тесноты связи. В данном случае это **индекс корреляции:**

$$\rho_{xy} = \sqrt{1 - \frac{\sigma^2_{\text{ост}}}{\sigma^2_y}},$$

где $\sigma^2_y = \frac{1}{n}\sum(y - \bar{y})^2$ – общая дисперсия результативного признака y,

$\sigma^2_{\text{ост}} = \frac{1}{n}\sum\left(y - y_x\right)^2$ – остаточная дисперсия.

на то, что нет необходимости усложнять форму уравнения регрессии и можно использовать линейную функцию.

Индекс детерминации используется для проверки существенности в целом уравнения регрессии по F-критерию Фишера: Величина данного показателя находится в пределах: $0 \le \rho_{xy} \le 1$. Чем ближе значение индекса корреляции к единице, тем теснее связь рассматриваемых признаков, тем более надежно уравнение регрессии.

Квадрат индекса корреляции носит название **индекса детерминации** и характеризует долю дисперсии результативного признака y, объясняемую регрессией, в общей дисперсии результативного признака:

$$\rho_{xy}^2 = 1 - \frac{\sigma_{\text{ост}}^2}{\sigma_y^2} = \frac{\sigma_{\text{объясн}}^2}{\sigma_y^2},$$

т.е. имеет тот же смысл, что и в линейной регрессии; $\sigma_{\text{объясн}}^2 = \frac{1}{n}\sum\left(y_x - \bar{y}\right)^2$.

Индекс детерминации ρ_{xy}^2 можно сравнивать с коэффициентом детерминации r_{xy}^2 для обоснования возможности применения линейной функции. Чем больше кривизна линии регрессии, тем величина r_{xy}^2 меньше ρ_{xy}^2. А близость этих показателей указывает

$$F = \frac{\rho_{xy}^2}{1-\rho_{xy}^2} \cdot \frac{n-m-1}{m},$$

где ρ_{xy}^2 – индекс детерминации,

n – число наблюдений,

m – число параметров при переменной x.

Фактическое значение F-критерия (1.23) сравнивается с табличным при уровне значимости α и числе степеней свободы $k_2 = n-m-1$ (для остаточной суммы квадратов) и $k_1 = m$ (для факторной суммы квадратов).

О качестве нелинейного уравнения регрессии можно также судить и по средней ошибке аппроксимации, которая, так же как и в линейном случае.

Пример на нелинейную модель . По семи предприятиям легкой промышленности региона получена информация, характеризующая зависимость объема выпуска продукции (*Y*, млн. руб.) от объема капиталовложений (*X*, млн. руб.).

Y	64	56	52	48	50	46	38
X	64	68	82	76	84	96	100

Требуется:

1 . Для характеристики *Y* от *X* построить следующие модели:

• линейную (для сравнения с нелинейными),

• степенную, показательную, гиперболическую.

2. Оценить каждую модель, определив:

• индекс корреляции,

• среднюю относительную ошибку,

• коэффициент детерминации,

• *F*-критерий Фишера.

3. Составить сводную таблицу вычислений, выбрать лучшую модель, дать интерпретацию рассчитанных характеристик.

4. Рассчитать прогнозные значения результативного признака по лучшей модели, если объем капиталовложений составит 89,573 млн. руб.

5. Результаты расчетов отобразить на графике.

Решение: **1. Построение линейной модели парной регрессии**

Определим линейный коэффициент парной корреляции по следующей формуле:

$$r_{Y,X} = \frac{\sum(y-\bar{y})\times(x-\bar{x})}{\sqrt{\sum(y-\bar{y})^2\times\sum(x-\bar{x})^2}} = \frac{-593{,}714}{\sqrt{397{,}71\times1077{,}71}} = -0{,}907.$$

Можно сказать, что связь между объемом капиталовложений *X* и объемом выпуска продукции *Y* обратная, достаточно сильная.

Уравнение линейной регрессии имеет вид: $\hat{y} = a + b\times x$.

Значения параметров *a* и *b* линейной модели определим, используя данные таблицы 7.

$$b = \frac{\overline{y\times x}-\bar{y}\times\bar{x}}{\overline{x^2}-\bar{x}^2} = \frac{4033{,}14-50{,}57\times81{,}43}{6784{,}57-81{,}43^2} = -0{,}55,$$

$$a = \bar{y} - b\times\bar{x} = 50{,}57 - 0{,}55\times81{,}43 = 95{,}36$$

Уравнение линейной регрессии имеет вид: $\hat{y} = 95{,}36 - 0{,}55\times x$.

Таблица 7. Расчет параметров линейной регрессии

t	*y*	*x*	$y\times x$	$x\times x$	$(y_i-\bar{y})$	$(y_i-\bar{y})^2$	$(x_i-\bar{x})$	$(x_i-\bar{x})^2$	$\hat{y}$	$e_i = y_i-\hat{y}$	$\left\|\frac{e_i}{y_i}\right\|\times100\%$
1	64	64	4096	4096	13,43	180,36	-17,4	303,8	60,2	3,84	6,000
2	56	68	3808	4624	5,43	29,485	-13,4	180,36	58,0	-1,96	-3,500
3	52	82	4264	6724	1,43	2,0449	0,57	0,3249	50,3	1,74	3,346
4	48	76	3648	5776	-2,57	6,6049	-5,43	29,485	53,6	-5,56	-11,583
5	50	84	4200	7056	-0,57	0,3249	2,57	6,6049	49,2	0,84	1 680
6	46	96	4416	9216	-4,57	20,885	14,57	212,28	42,6	3,44	7,478
7	38	100	3800	10000	-12,6	158.0	18,57	344,84	40,4	-2,36	-6,211
Итого	354,00	570,00	28232	47492	0,01	397,71		1077,7		-0,02	39,798
ср. знач.	50,57	81,43	4033,14	6784,57							5,685
диспер.	56,80	154,00									

С увеличением объема капиталовложений на 1 млн. руб. объем выпускаемой продукции уменьшится в среднем на 550 тыс. руб. Это свидетельствует о неэффективности работы предприятий, и необходимо принять меры для выяснения причин и устранения этого недостатка.

Рассчитаем коэффициент детерминации:

$$R^2 = r_{y-x}^2 = 0{,}822.$$

Вариация результата Y (объема выпуска продукции) на 82,2 % объясняется вариацией фактора X (объемом капиталовложений).

Оценку значимости уравнения регрессии проведем с помощью F-критерия Фишера:

$$F = \frac{r_{Y,X}^2}{1 - r_{Y,X}^2} \times (n-2) = \frac{0{,}822}{1 - 0{,}822} \times (7-2) = 23{,}09;$$

$F > F_{табл} = 6{,}61$ для $\alpha = 0{,}05$; $k_1 = m = 1$, $k_2 = n - m - 1 = 5$.

Уравнение регрессии с вероятностью 0,95 в целом статистически значимое, т. к. $F > F_{табл}$.

Определим среднюю относительную ошибку:

$$\overline{E}_{отн} = \frac{1}{n} \times \sum \frac{|E_i|}{y} \times 100\% = \frac{1}{n} \times \sum \left| \frac{y - \hat{y}}{y} \right| \times 100\% = \frac{39{,}798}{7} = 5{,}685\%.$$

В среднем расчетные значения $\hat{y}$ для линейной модели отличаются от фактчиеских значений на 5,685%.

2. Построение степенной модели парной регрессии

Уравнение степенной модели имеет вид: $\hat{y} = a \times x^b$.

Для построения этой модели необходимо произвести линеаризацию переменных. Для этого произведем логарифмирование обеих частей уравнения: $\lg \hat{y} = \lg a + b \lg x$.

Обозначим $Y = \lg y$, $X = \lg x$, $A = \lg a$.

Тогда уравнение примет вид: $Y = A + b X$ — линейное уравнение регрессии.

Рассчитаем его параметры, используя данные таблицы 8.

Таблица 8. Расчет параметров степенной регрессии

№	***y***	***Y=lgy***	***x***	***X =lgx***	***YX***	X^2	$\hat{y}$	e_i	**|e_i/y|×100%**	**ei^2**
1	64	1,8062	64	1,8062	3,2623	3,2623	61,294	2,706	4,23	7,322
2	56	1,7482	68	1,8325	3,2036	3,3581	58,066	-2,066	3,69	4,270
3	52	1,7160	82	1,9138	3,2841	3,6627	49,133	2,867	5,51	8,220
4	48	1,6812	76	1,8808	3,1621	3,5375	52,580	-4,580	9,54	20,976
5	50	1 ,6990	84	1,9243	3,2693	3,7029	48,088	1,912	3,82	3,657
6	46	1,6628	96	1,9823	3,2960	3,9294	42,686	3,314	7,20	10,982
7	38	1,5798	100	2,0000	3,1596	4,0000	41,159	-3,159	8,31	9,980
Итого	354	11,8931		13,3399	22,6370	25,4528		0,51	42,32	65,407

$$b = \frac{\overline{Y \cdot X} - \overline{Y} \cdot \overline{X}}{\overline{X^2} - \overline{X}^2} = \frac{3,2339 - 1,699 \cdot 1,9057}{3,6361 - 1,9057 \cdot 1,9057} = -0,8921,$$

$$A = \overline{Y} - b \cdot \overline{X} = 1,699 + 0,8921 \cdot 1,9057 = 3,3991.$$

Уравнение регрессии будет иметь вид: Y = 3,3991-0,8921 X.

Перейдем к исходным переменным x и y, выполнив потенцирование данного уравнения.

$$\hat{y} = 10^{3,399} \times x^{-0,892}$$

Получим уравнение степенной модели регрессии:

$$\hat{y} = 2506,915 \times x^{-0,892}.$$

Определим индекс корреляции:

$$\rho_{XY} = \sqrt{1 - \frac{\Sigma(y - \hat{y})^2}{\Sigma(y - \overline{y})^2}} = 0,914.$$

Связь между показателем y и фактором x можно считать достаточно сильной.

Коэффициент детерминации равен 0,836:

$$R^2 = \rho_{XY}^2 = 0,914^2 = 0,836.$$

Вариация результата Y (объема выпуска продукции) на 83,6% объясняется вариацией фактора X (объемом капиталовложений).

Рассчитаем F-критерий Фишера:

$$F = \frac{R^2}{1 - R^2} \times (n - 2) = \frac{0,836}{1 - 0,836} \times 5 = 25,5.$$

$F > F_{табл} = 6,61$ для $\alpha = 0,05$; $k_1 = m = 1$, $k_2 = n - m - 1 = 5$.

Уравнение регрессии с вероятностью 0,95 в целом статистически значимое, т.к. $F > F_{табл}$.

Средняя относительная ошибка

$$\overline{E}_{отн} = \frac{1}{n} \times \sum \frac{|E_i|}{y} \times 100\% = \frac{1}{n} \times \sum \left| \frac{y - \hat{y}}{y} \right| \times 100\%;$$

$$\overline{E}_{отн} = \frac{42,32}{7} = 6,04\%.$$

В среднем расчетные значения $\hat{y}$ для степенной модели отличаются от фактических значений на 6,04%.

3. Построение показательной функции

Уравнение показательной кривой: $\hat{y} = ab^x$.

Для построения этой модели необходимо произвести линеаризацию переменных. Для этого осуществим логарифмирование обеих частей уравнения: $\lg \hat{y} = \lg a + x \lg b$.

Обозначим: $Y = \lg \hat{y}$, $B = \lg b$, $A = \lg a$.

Получим линейное уравнение регрессии: $Y = A + B x$.

Рассчитаем его параметры, используя данные таблицы 9

Таблица 9. Расчет параметров показательной функции

t	***y***	***Y***	***x***	***Yx***	x^2	$Y-\overline{Y}$	$(Y-\overline{Y})^2$	$X-\overline{X}$	$(X-\overline{X}$	$\hat{y}$	$(y-\hat{y})^2$	e_i	**$\lvert e_i/y \rvert \times 100$**
1	64	1,8062	64	115,60	4096	0,1072	0,0115	-17,43	303,76	60,6	11,464	3,3859	5,290
2	56	1,7482	68	118,88	4624	0,0492	0,0024	-13,43	180,33	58	3,9632	-1,991	3,555
3	52	1,7160	82	140,71	6724	0,0170	0,0003	0,57	0,33	49,7	5,4221	2,3285	4,478
4	48	1,6812	76	127,77	5776	-0,017	0,0003	-5,43	29,47	53,1	25,804	-5,08	10,583
5	50	1,6990	84	142,71	7056	0,0000	0,0000	2,57	6,61	48,6	2,0031	1,4153	2,831
6	46	1,6628	96	159,62	9216	-0,036	0,0013	14,57	212,33	42,5	11,933	3,4544	7,509
7	38	1,5798	100	157,98	10000	-0,119	0,0142	18,57	344,90	40,7	7,3132	-2,704	7,117
итого	354	11,8931	570	963,28	4749		0,0300		1077,7		67,903	0,8093	41,363
Сред. знач.	50,57	1,6990	81,4	137,61	6785								5.909

$$b=\frac{\overline{Y\times x}-\overline{Y}\times\overline{x}}{\overline{x^2}-\overline{x}^2}=\frac{137{,}61-1{,}699\cdot 81{,}43}{6785-81{,}43\cdot 81{,}43}=-0{,}0048;$$

$$A=\overline{Y}-b\times\overline{x}=1{,}699+0{,}0048\cdot 81{,}43=2{,}09.$$

Уравнение будет иметь вид: Y = 2,09 + 0,0048 x.

Перейдем к исходным переменным x и y, выполнив потенцирование данного уравнения:

$$\hat{y}=10^{2{,}09}\times(10^{-0{,}0048})^x=123{,}03\times 0{,}989^x.$$

Определим индекс корреляции:

$$\rho_{XY}=\sqrt{1-\frac{\Sigma(y-\hat{y})^2}{\Sigma(y-\overline{y})^2}}=\sqrt{1-\frac{67{,}9}{397{,}71}}=0{,}91.$$

вязь между показателем y и фактором x можно считать тесной.

Индекс детерминации: $R^2=\rho_{XY}^2=0{,}91^2=0{,}828.$

Вариация результата Y (объема выпуска продукции) на 41,1 % объясняется вариацией фактора X (объем капиталовложений).

Рассчитаем F-критерий Фишера:

$$F = \frac{R^2}{1-R^2} \times (n-2) = 24{,}06.$$

$F > F_{табл} = 6{,}61$ для $\alpha = 0{,}05$; $k_1 = m = 1$, $k_2 = n - m - 1 = 5$.

Уравнение регрессии с вероятностью 0,95 в целом статистически значимое, т. к. $F > F_{табл}$.

Средняя относительная ошибка: $\bar{E}_{отн} = \frac{41{,}363}{7} = 5{,}909\%.$

В среднем расчетные значения $\hat{y}$ для линейной модели отличаются от фактических значений на 5,909 %.

4. Построение гиперболической функции

Уравнение гиперболической функции:

Произведем линеаризацию модели путем замены $X = 1/x$. В результате получим линейное уравнение $\hat{y} = a + b / x$.

Рассчитаем его параметры по данным таблицы 10.

$$A = \bar{Y} - b \times \bar{x} = 50{,}57 - 3571{,}9 \cdot 0{,}0126 = 5{,}7.$$

$$b = \frac{\overline{y \times X} - \bar{y} \times \bar{X}}{\overline{X^2} - \bar{X}^2} = \frac{0{,}6941 - 50{,}57 \cdot 0{,}0126}{0{,}0001618 - 0{,}0126 \cdot 0{,}0126} = 3571{,}95;$$

Получим следующее уравнение гиперболической модели: $\hat{y} = 5{,}7 - 3571{,}9 / x.$

Определим индекс корреляции:

$$\rho_{YX} = \sqrt{1 - \frac{\Sigma(y-\hat{y})^2}{\Sigma(y-\bar{y})^2}} = \sqrt{1 - \frac{65{,}159}{397{,}71}} = 0{,}914.$$

Связь между показателем y и фактором x можно считать достаточно сильной.

Индекс детерминации: $R^2 = \rho_{YX}^2 = 0{,}914^2 = 0{,}835.$

Вариация результата Y (объема выпуска продукции) на 83,5% объясняется вариацией фактора X (объемом капиталовложений).

Таблица 10. Расчет параметров гиперболической функции

t	***y***	***x***	***X***	***yX***	X^2	$Y - \bar{Y}$	$(Y - \bar{Y})$	$(y - \hat{y})$	e_i	$\hat{y}$	$\lvert e_i/y \rvert \times 100$
1	64	64	0,0156	1,0000	0,0002441	13,43	180,33	61,5	2,489	6,195	3,889
2	56	68	0,0147	0,8235	0,0002163	5,43	29,47	58,2	-2,228	4,963	3,978
3	52	82	0,0122	0,6341	0,0001487	1,43	2,04	49,3	2,740	7,508	5,270

4	48	76	0,0132	0,6316	0,0001731	-2,57	6,61	52,7	-4,699	22,07	9,789
5	50	84	0,0119	0,5952	0,0001417	-0,57	0,32653	48,2	1,777	3,159	3,555
6	46	96	0,0104	0,4792	0,0001085	-4,57	20,90	42.9	3,093	9,564	6,723
7	38	100	0,0100	0,3800	0,0001000	-12,57	158.04	41,4	-3,419	11,69	8,997
итого	354		0,0880	4,5437	0,0011325		397,71	354,2	-0,246	65,15	42,202
Сред. знач.	50,5		0,0126	0,6491	0,0001618						6,029

F-критерий Фишера:

$$F = \frac{R^2}{1-R^2} \times (n-2) = \frac{0,835}{1-0,835} \times 5 = 25,3.$$

$F > F_{табл} = 6,61$ для $\alpha = 0,05$; $k_1 = m = 1$, $k_2 = n - m - 1 = 5$.

Уравнение регрессии с вероятностью 0,95 в целом не является статистически значимым, т. к. $F > F_{табл}$.

Определим среднюю относительную ошибку:

$$\overline{E}_{отн} = \frac{1}{n} \sum \frac{|E_i|}{y} \times 100\% = \frac{42,202}{7} = 6,029\%.$$

В среднем расчетные значения для линейной модели отличаются от фактических значений на 6,029 %. Для выбора лучшей модели построим сводную таблицу результатов (таблица 11).

Таблица 11 Сравнение результатов

Параметры Модель	**Коэффициент детерминации R^2**	**F-критерий Фишера**	**Индекс корреляции $\rho_{YX}(r_{YX})$**	**Средняя относительная ошибка аппроксимации**
Линейная	0,822	23,09	0,907	5,685
Степенная	0,828	24,06	0,910	6,054
Показательная	0,828	24,06	0,910	5,909
Гиперболическая	0,835	25,30	0,914	6,029

Все модели имеют примерно одинаковые характеристики, но большее значение F-критерия Фишера и большее значение коэффициента детерминации R^2 имеет гиперболическая модель. Ее можно взять в качестве лучшей для построения прогноза.

Расчет прогнозного значения результативного показателя:

Прогнозное значение результативного признака (объема выпуска продукции) определим по уравнению гиперболической модели, подставив в него планируемую (заданную по условию) величину объема капиталовложений:

$$\hat{y}_{пр} = 5{,}7 + 3571{,}9 \times X_{пр} = 5{,}7 + 3571{,}9 / 89{,}573 = 45{,}542 \ (\text{млн. руб.})$$

3.3 Адаптивная модель Брауна

Пример. Рассчитайте экспоненциальную среднюю для временного ряда индекса потребительских цен в Кемеровской области (таблица 12). В качестве начального значения экспоненциальной средней возьмите среднее значение из 5 первых уровней ряда. Расчеты проведите для двух различных значений параметров адаптации α: α = 0,1; α = 0,5.

Таблица 12.Индекс потребительских цен Кемеровской области

t	y_t	t	y_t	t	y_t	t	y_t
1	102,1	10	100,7	19	101,3	28	100,6
2	100,7	11	100,5	20	100,6	29	100,4
3	100,8	12	100,6	21	100,8	30	100,2
4	100,3	13	100,3	22	100,7	31	100,4
5	100,6	14	100,3	23	100,6	32	100,3
6	99,9	15	100,6	24	100,5	33	100,1
7	100,3	16	100,3	25	101	34	101,2
8	99,9	17	100,4	26	100,8		
9	100,1	18	100,8	27	100,5		

Решение. Определим нулевую среднюю

$$S_0 = \frac{102{,}1 + 100{,}7 + 100{,}8 + 100{,}3 + 100{,}6}{5} = 100{,}9 .$$

Найдем значения экспоненциальной средней при α = 0,1.

$$S_1 = \alpha y_1 + (1-\alpha) S_0 = 0{,}1 \cdot 102{,}1 + 0{,}9 \cdot 100{,}9 = 101{,}02 ,$$

$$S_2 = \alpha y_2 + (1-\alpha) S_1 = 0{,}1 \cdot 100{,}7 + 0{,}9 \cdot 101{,}2 = 100{,}99$$

и т.д.

Результаты расчетов представлены в таблице 13.

Проведем аналогичные расчеты для α = 0,5.

$$S_1 = \alpha y_1 + (1-\alpha) S_0 = 0{,}5 \cdot 102{,}1 + 0{,}5 \cdot 100{,}9 = 101{,}5 ,$$

$$S_2 = \alpha y_2 + (1-\alpha) S_1 = 0{,}5 \cdot 100{,}7 + 0{,}5 \cdot 101{,}5 = 101{,}1$$

и т.д.

Результаты расчетов представлены в таблице 13.

На рисунке 3 наглядно представлено влияние значения параметра адаптации α на характер сглаженного ряда.

При α = 0,1 экспоненциальная средняя носит более гладкий характер, т. к. в этом случае в наибольшей степени поглощаются случайные колебания временного ряда.

Таблица 13. Экспоненциальные средние

t	Экспоненциальная средняя		t	Экспоненциальная средняя	
	$\alpha = 0{,}1$	$\alpha = 0{,}5$		$\alpha = 0{,}1$	$\alpha = 0{,}5$
1	101,02	101,50	18	100,54	100,60
2	100,99	101,10	19	100,61	100,95
3	100,97	100,95	20	100,61	100,77
4	100,90	100,63	21	100,63	100,79
5	100,87	100,61	22	100,64	100,74
6	100,77	100,26	23	100,63	100,67
7	100,73	100,28	24	100,62	100,59
8	100,64	100,09	25	100,66	100,79
9	100,59	100,09	26	100,67	100,80
10	100,60	100,40	27	100,65	100,65
11	100,59	100,45	28	100,65	100,62
12	100,59	100,52	29	100,62	100,51
13	100,56	100,41	30	100,58	100,36
14	100,54	100,36	31	100,56	100,38
15	100,54	100,48	32	100,54	100,34
16	100,52	100,39	33	100,49	100,22
17	100,51	100,39	34	100,56	100,71

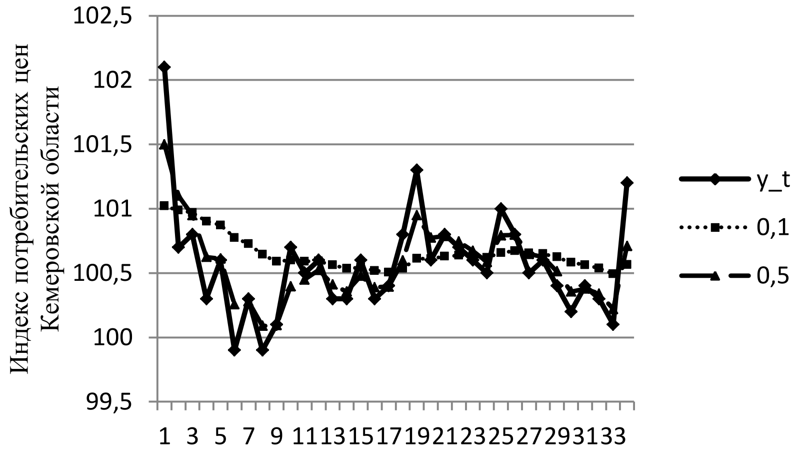

Рис. 3. Экспоненциальное сглаживание временного ряда индекса потребительских цен Кемеровской области

Можно разобрать несколько моделей для разных значений α и выбрать лучшую на основе минимума среднеквадратической ошибки модели.

3.4 Моделирование сезонности

Пример 1. Пусть имеются некоторые условные данные об общем количестве правонарушений на таможне одного из субъектов РФ (таблица 14).

Таблица 14 Данные для анализа

Год	Квартал	t	Количество возбуж-денных дел, y_t
2014	I	1	375
	II	2	371
	III	3	869
	IV	4	1015
2015	I	5	357
	II	6	471
	III	7	992
	IV	8	1020
2016	I	9	390
	II	10	355
	III	11	992
	IV	12	905
2017	I	13	461
	II	14	454
	III	15	920
	IV	16	927

Построим поле корреляции (рисунок 4):

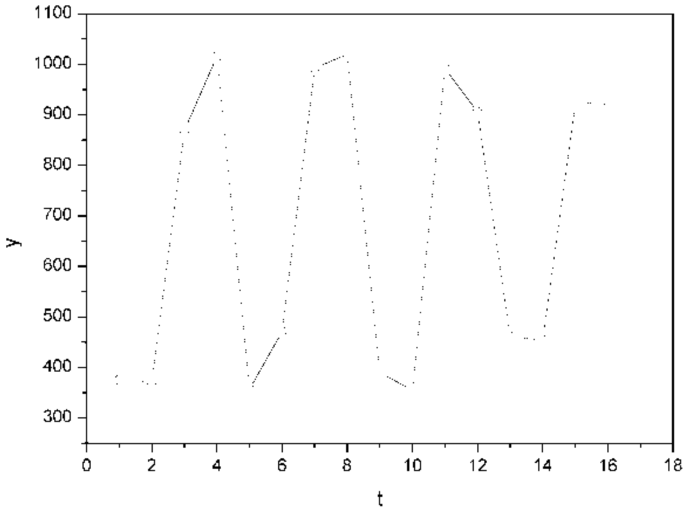

Рис. 4 – Корреляционное поле

Исходя из графика видно, что значения y образуют пилообразную фигуру. Рассчитаем несколько последовательных коэффициентов автокорреляции. Для этого составляем первую вспомогательную таблицу 15.

Таблица 15 Расчет коэффициента автокорреляции 1-го порядка

t	y_t	y_{t-1}	$y_t-\overline{y}_1$	$y_{t-1}-\overline{y}_2$	$\left(y_t-\overline{y}_1\right)\times\left(y_{t-1}-\overline{y}_2\right)$	$\left(y_t-\overline{y}_1\right)^2$	$\left(y_{t-1}-\overline{y}_2\right)^2$
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	375	–	–	–	–	–	–
2	371	375	-328,33	-288,13	94601,72	107800,59	83018,90
3	869	371	169,67	-292,13	-49565,70	28787,91	85339,94
4	1015	869	315,67	205,87	64986,98	99647,55	42382,46
5	357	1015	-342,33	351,87	-120455,66	117189,83	123812,50
6	471	357	-228,33	-306,13	69898,66	52134,59	93715,58
7	992	471	292,67	-192,13	-56230,69	85655,73	36913,94
8	1020	992	320,67	328,87	105458,74	102829,25	108155,48
9	390	1020	-309,33	356,87	-110390,60	95685,05	127356,20
10	355	390	-344,33	-273,13	94046,85	118563,15	74600,00
11	992	355	292,67	-308,13	-90180,41	85655,73	94944,10
12	905	992	205,67	328,87	67638,69	42300,15	108155,48
13	461	905	-238,33	241,87	-57644,88	56801,19	58501,10
14	454	461	-245,33	-202,13	49588,55	60186,81	40856,54
15	920	454	220,67	-209,13	-46148,72	48695,25	43735,36
16	927	920	227,67	256,87	58481,59	51833,63	65982,20
Сумма	10499	9947	9,05	0,05	74085,16	1153766,39	1187469,73
Среднее значение	699,33	663,13	–	–	–	–	–

Следует заметить, что среднее значение получается путем деления не на 16, а на 15, т.к. у нас теперь на одно наблюдение меньше.

Теперь вычисляем коэффициент автокорреляции первого порядка по формуле:

$$r_1=\frac{74085,16}{\sqrt{1153766,39\cdot 1187469,73}}=0,063294\cdot$$

Аналогично находим коэффициенты автокорреляции более высоких порядков, а все полученные значения заносим в сводную таблицу 16.

Таблица 16 коэффициенты автокорреляции

Лаг	Коэффициент автокорреляции уровней
1	0,063294
2	–0,961183
3	–0,036290
4	0,964735
5	0,050594
6	–0,976516
7	–0,069444
8	0,964629

9	0,162064
10	-0,972918
11	-0,065323
12	0,985761

Анализ коррелограммы и графика исходных уровней временного ряда позволяет сделать вывод о наличии в изучаемом временном ряде сезонных колебаний периодичностью в четыре квартала. Построим аддитивную модель временного ряда. Для этого проведем ряд шагов.

Шаг 1. Проведем выравнивание исходных уровней ряда методом скользящей средней. Для этого:

1.1. Просуммируем уровни ряда последовательно за каждые четыре квартала со сдвигом на один момент времени и определим условные годовые объемы потребления электроэнергии (гр. 3 табл. 17).

1.2. Разделив полученные суммы на 4, найдем скользящие средние (гр. 4 табл. 17). Полученные таким образом выровненные значения уже не содержат сезонной компоненты.

1.3. Приведем эти значения в соответствие с фактическими моментами времени, для чего найдем средние значения из двух последовательных скользящих средних – центрированные скользящие средние (гр. 5 табл. 17).

Шаг 2. Найдем оценки сезонной компоненты как разность между фактическими уровнями ряда и центрированными скользящими средними (гр. 6 табл. 17). Используем эти оценки для расчета значений сезонной компоненты S (табл. 18). Для этого найдем средние за каждый квартал (по всем годам) оценки сезонной компоненты S_i. В моделях с сезонной компонентой обычно предполагается, что сезонные воздействия за период взаимопогашаются.

Таблица 17. Расчет скользящих средних и оценка сезонной компоненты

№ квартала, t	Количество правонарушений, y_t	Итого за четыре квартала	Скользящая средняя за четыре квартала	Центрированная скользящая средняя	Оценка сезонной компоненты
1	**2**	**3**	**4**	**5**	**6**
1	375	–	–	–	–
2	371	2630	657,5	–	–
3	869	2612	653	655,25	213,75
4	1015	2712	678	665,5	349,5
5	357	2835	708,75	693,75	-336,75
6	471	2840	710	709,375	-238,375
7	992	2873	718,25	714,125	277,875
8	1020	2757	689,25	703,75	316,25
9	390	2757	689,25	689,25	-299,25
10	355	2642	660,5	674,875	-319,875
11	992	2713	678,25	669,375	322,625
12	905	2812	703	690,625	214,375
13	461	2740	685	694	-233
14	454	2762	690,5	687,75	-233,75
15	920	–	–	–	–
16	927	–	–	–	–

В аддитивной модели это выражается в том, что сумма значений сезонной компоненты по всем кварталам должна быть равна нулю (таблица 18).

Таблица 18. Оценка скорректированной сезонной компоненты

Показатели	Год	№ квартала, i			
		I	II	III	IV
	2012	–	–	213,75	349,5
	2013	-336,75	-238,375	277,875	316,25
	2014	-299,25	-319,875	322,625	214,375
	2015	-233	-233,75	–	–
Всего за i -й квартал		-869	-792	814,25	880,125
Средняя оценка сезонной компоненты для i -го квартала, $\bar{S}_i$		-289,667	-264	271,417	293,375
Скорректированная сезонная компонента, S_i		-292,448	-266,781	268,636	290,593

Для данной модели имеем:

$-289,667-264+271,417+293,375=11,125$.

Корректирующий коэффициент: $k=11,125/4=2,781$.

Рассчитываем скорректированные значения сезонной компоненты ($S_i=\bar{S}_i-k$) и заносим полученные данные в таблицу 5.6.

Проверим равенство нулю суммы значений сезонной компоненты:

$$-292,448-266,781+268,636+290,593=0.$$

Шаг 3. Исключим влияние сезонной компоненты, вычитая ее значение из каждого уровня исходного временного ряда. Получим величины $T+E=Y-S$ (гр. 4 табл. 19). Эти значения рассчитываются за каждый момент времени и содержат только тенденцию и случайную компоненту (табл.19).

Таблица 19. Расчет параметров аддитивной модели

t	y_t	S_i	y_t-S_i	T	$T+S$	$E=y_t-(T+S)$	E^2
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	375	-292,448	667,448	672,700	380,252	-5,252	27,584
2	371	-266,781	637,781	673,624	406,843	-35,843	1284,721
3	869	268,636	600,364	674,547	943,183	-74,183	5503,117
4	1015	290,593	724,407	675,470	966,063	48,937	2394,830
5	357	-292,448	649,448	676,394	383,946	-26,946	726,087
6	471	-266,781	737,781	677,317	410,536	60,464	3655,895
7	992	268,636	723,364	678,240	946,876	45,124	2036,175
8	1020	290,593	729,407	679,163	969,756	50,244	2524,460

9	390	-292,448	682,448	680,087	387,639	2,361	5,574
10	355	-266,781	621,781	681,010	414,229	-59,229	3508,074
11	992	268,636	723,364	681,933	950,569	41,431	1716,528
12	905	290,593	614,407	682,857	973,450	-68,450	4685,403
13	461	-292,448	753,448	683,780	391,332	69,668	4853,630
14	454	-266,781	720,781	684,703	417,922	36,078	1301,622
15	920	268,636	651,364	685,627	954,263	-34,263	1173,953
16	927	290,593	636,407	686,550	977,143	-50,143	2514,320

Шаг 4. Определим компоненту T данной модели. Для этого проведем аналитическое выравнивание ряда ($T+E$) с помощью линейного тренда. Результаты аналитического выравнивания следующие:

$$T = 671,777 + 0,9233 \cdot t \cdot$$

Подставляя в это уравнение значения $t = 1, 2, ..., 16$, найдем уровни T для каждого момента времени (гр. 5 табл. 19).

Шаг 5. Найдем значения уровней ряда, полученные по аддитивной модели. Для этого прибавим к уровням T значения сезонной компоненты для соответствующих кварталов (гр. 6 табл. 19).

На одном графике (рис. 5) отложим фактические значения уровней временного ряда и теоретические, полученные по аддитивной модели.

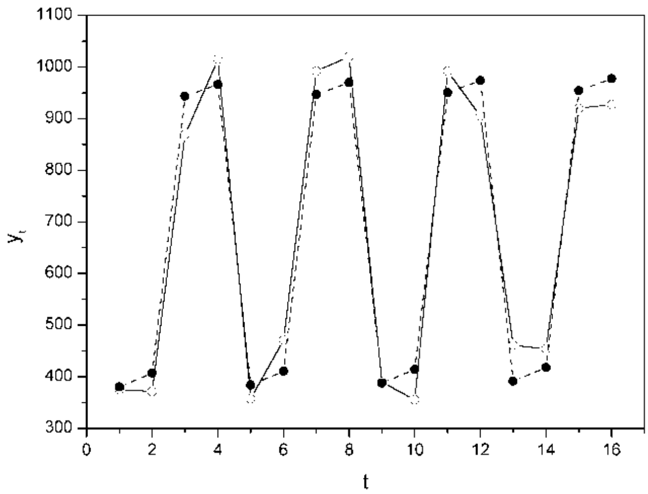

Рис. 5. – Исходные данные и аддитивная модель.

Для оценки качества построенной модели применим сумму квадратов полученных абсолютных ошибок.

$$R^2 = 1 - \frac{E^2}{(y_t - \bar{y})^2} = 1 - \frac{37911,973}{1252743,75} = 0,970 \cdot$$

Следовательно, можно сказать, что аддитивная модель объясняет 97% общей вариации уровней временного ряда количества правонарушений по кварталам за 4 года.

Шаг 6. Прогнозирование по аддитивной модели. Предположим, что по нашему примеру необходимо дать прогноз об общем объеме правонарушений на I и II кварталы 2018 года. Прогнозное значение F_t уровня временного ряда в аддитивной модели есть сумма трендовой и сезонной компонент. Для определения трендовой компоненты воспользуемся уравнением тренда

$$T = 671,777 + 0,9233 \cdot t.$$

Получим

$$T_{17} = 671,777 + 0,9233 \cdot 17 = 687,473;$$

$$T_{18} = 671,777 + 0,9233 \cdot 18 = 688,396.$$

Значения сезонных компонент за соответствующие кварталы равны: $S_1 = -292,448$ и $S_2 = -266,781$. Таким образом,

$$F_{17} = T_{17} + S_1 = 687,473 - 292,448 \approx 395;$$

$$F_{18} = T_{18} + S_2 = 688,396 - 266,781 \approx 422.$$

Т.е. в первые два квартала 2018 г. следовало ожидать порядка 395 и 422 правонарушений соответственно.

СПИСОК ЛИТЕРАТУРЫ

Основная учебная литература

1. Галочкин, В. Т. Эконометрика : учебник и практикум для вузов / В. Т. Галочкин. — Москва : Издательство Юрайт, 2020. — 288 с. — (Высшее образование). — ISBN 978-5-534-10751-7. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: https://urait.ru/bcode/467904
2. Эконометрика : учебник для вузов / И. И. Елисеева [и др.]. — Москва : Издательство Юрайт, 2020. — 449 с. — (Высшее образование). — ISBN 978-5-534-00313-0. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: https://urait.ru/bcode/449677 .

Дополнительная литература

3. Костюнин, В. И. Эконометрика : учебник и практикум для вузов / В. И. Костюнин. — Москва : Издательство Юрайт, 2020. — 285 с. — (Высшее образование). — ISBN 978-5-534-02660-3. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: https://urait.ru/bcode/450113
4. Бородич, С. А. Эконометрика. Практикум : учеб. пособие / С.А. Бородич. — Минск : Новое знание ; Москва : ИНФРА-М, 2018. — 329 с. : ил. — (Высшее образование: Бакалавриат). - ISBN 978-5-16-100513-2. - Текст : электронный. - URL: https://new.znanium.com/catalog/product/988809
5. Яковлев, В. П. Эконометрика : учебник для бакалавров / В. П. Яковлев. — Москва : Издательско-торговая корпорация «Дашков и К°», 2019. - 384 с. - ISBN 978-5-394-02532-7. - Текст : электронный. - URL: https://znanium.com/catalog/product/1091204 – Режим доступа: по подписке.
6. Мардас, А. Н. Эконометрика : учебник и практикум для вузов / А. Н. Мардас. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2020. — 180 с. — (Высшее образование). — ISBN 978-5-9916-8164-3. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: https://urait.ru/bcode/451492.
7. Подкорытова, О. А. Анализ временных рядов : учебное пособие для вузов / О. А. Подкорытова, М. В. Соколов. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2020. — 267 с. — (Высшее образование). — ISBN 978-5-534-02556-9. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: https://urait.ru/bcode/450587 .
8. Зелепухин, Ю.В. Эконометрика : учебно-методическое пособие : [12+] / Ю.В. Зелепухин. – Москва ; Берлин : Директ-Медиа, 2020. – 123 с. : табл., ил. – Режим доступа: по подписке. – URL: http://biblioclub.ru/index.php?page=book&id=572682 – Библиогр.: с. 92. – ISBN 978-5-4499-0573-4. – DOI 10.23681/572682. – Текст : электронный.

Приложение 1- Пример титульного листа

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Новокузнецкий институт (филиал)
федерального государственного бюджетного образовательного
учреждения высшего образования
«Кемеровский государственный университет»

Факультет информатики, математики и экономики

Кафедра экономики и управления

Иванов Иван Иванович

гр. ЭФКаз-18

Контрольная работа

по дисциплине «ЭКОНОМЕТРИКА»

ВАРИАНТ №

по специальности 38.05.01 Экономическая безопасность

Проверил:

канд. экон. наук, доцент

М.А. Кречетова

Общий балл:________________

Оценка: ____________________

____

подпись

«___» ________ 2020 г.

Новокузнецк, 2020