Министерство образования и науки Российской Федерации

ФГБОУ ВПО «Российский химико–технологический университет им. Д.И. Менделеева»

Новомосковский институт (филиал)

кафедра «Экономика, финансы и бухгалтерский учет»

**Е. Д. Саяпина**

**Эконометрика**

**методические указания по изучению курса**

**и задания на курсовую работу**

**для студентов заочного отделения, обучающихся по направлению**

**080100 62 «Экономика»**

Новомосковск

2014

Содержание

[1 Организационно-методические указания 4](#_Toc409463501)

[2 Структура и содержание дисциплины 5](#_Toc409463502)

[3 Требования, предъявляемые к выполнению и защите курсовой работы 9](#_Toc409463503)

[4 Методические указания по выполнению курсовой работы 10](#_Toc409463504)

[4.1 Исследование взаимосвязи между переменными 10](#_Toc409463505)

[4.2 Построение уравнения парной линейной регрессии 13](#_Toc409463506)

[4.3 Проверка качества регрессионной модели 14](#_Toc409463507)

[4.3.1 Проверка адекватности уравнения регрессии эмпирическим данным 14](#_Toc409463508)

[4.3.2 Проверка адекватности уравнения регрессии данным генеральной совокупности 17](#_Toc409463509)

[4.3.3 Проверка статистической значимости коэффициентов регрессии 19](#_Toc409463510)

[4.4 Проверка свойств случайного отклонения 22](#_Toc409463511)

[4.4.1 Автокорреляция остатков 23](#_Toc409463512)

[4.4.2 Гомо- и гетероскедастичность 27](#_Toc409463513)

[4.5 Изменение спецификации модели 34](#_Toc409463514)

[4.5.1. Спецификация регрессионных моделей 34](#_Toc409463515)

[4.5.2 Ошибки спецификации 36](#_Toc409463516)

[4.5.3 Построение множественной регрессионной модели 38](#_Toc409463517)

[4.5.4 Мультиколлинеарность 39](#_Toc409463518)

[4.6 Сравнение качества моделей 39](#_Toc409463519)

[5 Задания на курсовую работу 42](#_Toc409463520)

[Вариант 1 42](#_Toc409463521)

[Вариант 2 44](#_Toc409463522)

[Вариант 3 46](#_Toc409463523)

[Вариант 4 48](#_Toc409463524)

[Вариант 5 50](#_Toc409463525)

[Вариант 6 52](#_Toc409463526)

[Вариант 7 54](#_Toc409463527)

[Вариант 8 56](#_Toc409463528)

[Вариант 9 58](#_Toc409463529)

[Вариант 10 60](#_Toc409463530)

[Вариант 11 62](#_Toc409463531)

[Вариант 12 64](#_Toc409463532)

[Вариант 13 66](#_Toc409463533)

[Вариант 14 68](#_Toc409463534)

[Вариант 15 70](#_Toc409463535)

[Вариант 16 72](#_Toc409463536)

[Вариант 17 74](#_Toc409463537)

[Вариант 18 76](#_Toc409463538)

[Вариант 19 78](#_Toc409463539)

[Вариант 20 80](#_Toc409463540)

[Вариант 21 82](#_Toc409463541)

[Вариант 22 84](#_Toc409463542)

[Вариант 23 86](#_Toc409463543)

[Вариант 24 88](#_Toc409463544)

[Вариант 25 90](#_Toc409463545)

[Вариант 26 92](#_Toc409463546)

[Вариант 27 94](#_Toc409463547)

[Вариант 28 96](#_Toc409463548)

[Вариант 29 98](#_Toc409463549)

[Рекомендуемая литература 100](#_Toc409463550)

[Приложение А 101](#_Toc409463551)

[Приложение Б 102](#_Toc409463552)

[Приложение В 103](#_Toc409463553)

# 1 Организационно-методические указания

Дисциплина «Эконометрика» является дисциплиной базовой части профессионального цикла ФГОС ВПО по направлению 080100 62 «Экономика».

Изучение дисциплины «Эконометрика» основывается на базе знаний, умений и компетенций, полученных студентами в ходе освоения дисциплин «Математический анализ», «Теория вероятностей и математическая статистика», «Поиск и обработка экономической информации» математического и естественнонаучного цикла, «Микроэкономика» и «Макроэкономика» профессионального цикла.

Целью освоения дисциплины «Эконометрика» является получение базовых знаний и формирование основных навыков в области стохастического моделирования и анализа деятельности финансово-экономических объектов.

В результате изучения дисциплины студент должен:

**- знать:**

- основные термины и понятия, связанные с построением и анализом эконометрических моделей;

- основные виды стохастических моделей, используемые для моделирования финансово-экономических объектов;

- теоретические основы вероятностно-статистических методов математического моделирования и проверки качества математических моделей;

**- уметь:**

- производить спецификацию финансово-экономических моделей;

- применять математический аппарат регрессионного анализа для оценки параметров финансово-экономических моделей;

- проверять качество эконометрических моделей и анализировать их параметры с помощью аппарата статистической проверки гипотез;

- использовать построенные эконометрические модели финансово-экономических объектов для прогнозирования экономической ситуации и выработки рекомендаций;

**- владеть:**

- навыками построения финансово-экономических моделей и их анализа;

- навыками применения прикладных программных средств для анализа экономических данных.

Студенты заочного отделения самостоятельную подготовку по курсу «Эконометрика» должны начинать:

* с изучения теоретического материала, изложенного в рекомендованных учебниках, учебных пособиях;
* с изучения способов решения практических задач, разобранных в рекомендованных учебниках, задачниках, практикумах.

После изучения вышеуказанных материалов студенты могут приступать к выполнению курсовой работы.

**Формы контроля:**

- выполнение и защита курсовой работы;

- зачет, включающий в себя проверку знания теоретических вопросов и умения решать практические задачи.

# 2 Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 4 зачетных единицы, 144 часа.

Таблица 1 – Распределение часов по темам и видам занятий:

|  |  |
| --- | --- |
| Наименование темы | Количество часов |
| лекции | практические занятия | лабораторные работы | самостоятельная работа студента |
| Установочная лекция | 1 | - | - | - |
| 1 Эконометрика как наука | 0,5 | - | - | 4 |
| 2 Линейная модель парной регрессии | 1,5 | 0,5 | 2 | 23 |
| 3 Линейная модель множественной регрессии | 0,5 | 0,5 | 1 | 15 |
| 4 Нелинейные модели регрессии и их линеаризация | 0,5 | 0,5 | 1 | 15 |
| 5 Регрессионные модели с гетероскедастичными и автокоррелированными остатками | 0,5 | 0,5 | - | 20 |
| 6 Регрессионные модели с переменной структурой (фиктивные переменные) | - | 1 | - | 15 |
| 7 Временные ряды | 1 | 1 | 2 | 17 |
| 8 Системы линейных одновременных уравнений | 0,5 | - | - | 15 |
| **Всего** | **6** | **4** | **6** | **124** |

***Тема 1. Эконометрика как наука***

Задачи эконометрики. Связь эконометрики с математикой, статистикой, экономической теорией. Экономические модели и принципы их спецификации. Корреляционная и регрессионная зависимость между экономическими переменными.

*Вопросы для самоконтроля:*

1. Назовите основные задачи эконометрики.
2. В чем проявляется связь эконометрики с математикой, статистикой, экономической теорией?
3. Перечислите основные типы экономических моделей.
4. Перечислите способы обнаружения корреляционной зависимости между экономическими переменными.
5. Что показывает функция регрессии?
6. Что показывает регрессионная модель?
7. Назовите причины присутствия в регрессионных моделях случайного фактора.

***Тема 2. Линейная модель парной регрессии***

Оценивание параметров линейной модели методом наименьших квадратов (МНК). Статистические свойства МНК-оценок параметров регрессии. Показатели качества регрессии. Проверка гипотез относительно коэффициентов линейного уравнения регрессии. Проверка адекватности уравнения регрессии. Прогнозирование зависимой переменной.

*Вопросы для самоконтроля:*

1. Дайте определение парной линейной регрессии.
2. В чем отличие теоретического и эмпирического уравнений регрессии?
3. В чем суть метода наименьших квадратов?
4. Как интерпретируются параметры парного линейного уравнения регрессии?
5. Перечислите предпосылки МНК. С чем связана необходимость их выполнения?
6. Как рассчитывается стандартная ошибка регрессии? Как ее интерпретировать?
7. Как рассчитываются стандартные ошибки коэффициентов регрессии? Поясните их интерпретацию.
8. Сформулируйте гипотезу о равенстве одного из коэффициентов регрессии некоторому ожидаемому значению и проверьте ее с помощью аппарата статистической проверки гипотез. Какой критерий вы при этом использовали?
9. Сформулируйте гипотезу о статистической значимости одного из коэффициентов регрессии и проверьте ее с помощью аппарата статистической проверки гипотез. Какой критерий вы при этом использовали?
10. Что показывает доверительный интервал для коэффициента регрессии?
11. Что показывает коэффициент детерминации? В каких пределах он изменяется?
12. Как осуществляется прогнозирование среднего значения зависимой переменной?
13. Что показывает доверительный интервал для среднего значения зависимой переменной?

***Тема 3. Линейная модель множественной регрессии***

Оценивание параметров множественной регрессии. Статистические характеристики оценок коэффициентов множественной регрессии. Интервальные оценки коэффициентов множественной регрессии. Проверка статистической значимости коэффициентов эмпирического уравнения регрессии. Проверка общего качества эмпирического уравнения регрессии.

*Вопросы для самоконтроля:*

1. Что показывают коэффициенты множественной регрессии?
2. Какое соотношение должно выполняться между количеством переменных в линейной регрессионной модели и числом наблюдений?
3. Перечислите предпосылки МНК для множественной регрессии.
4. Запишите выражение для вычисления коэффициентов регрессии в матричной форме.
5. Как рассчитывается стандартная ошибка множественной регрессии? Как ее интерпретировать?
6. Как рассчитываются стандартные ошибки коэффициентов регрессии в случае применения матричной формы МНК?

***Тема 4. Нелинейные модели регрессии и их линеаризация***

Логарифмические модели. Гиперболические модели. Степенные модели. Показательные и экспоненциальные модели. Полиномиальные модели. Оценка качества моделей нелинейной регрессии.

*Вопросы для самоконтроля:*

1. Для описания каких экономических объектов применяются логарифмические модели? Как проводится линеаризация логарифмической модели?
2. Для описания каких экономических объектов применяются гиперболические модели? Как проводится линеаризация гиперболической модели?
3. Для описания каких экономических объектов применяются степенные модели? Как проводится линеаризация степенной модели?
4. Для описания каких экономических объектов применяются показательные и экспоненциальные модели? Как проводится линеаризация показательной и экспоненциальной модели?
5. Для описания каких экономических объектов применяются полиномиальные модели? Как проводится линеаризация полиномиальной модели?
6. Назовите особенности оценки качества нелинейной регрессии.

***Тема 5. Регрессионные модели с гетероскедастичными и автокоррелированными остатками***

Ошибки спецификации. Рекомендации по обнаружению и исправлению ошибок спецификации. Гетероскедастичность (суть, последствия, методы обнаружения и устранения). Автокорреляция остатков (суть, причины, последствия, методы обнаружения и устранения). Обобщенный метод наименьших квадратов.

*Вопросы для самоконтроля:*

1. Назовите основные показатели качества эконометрической модели.
2. Назовите основные ошибки спецификации.
3. Как можно обнаружить ошибку спецификации?
4. В чем суть гетероскедастичности?
5. Каковы последствия гетероскедастичности?
6. Опишите методы обнаружения гетероскедастичности.
7. Дайте определение автокорреляции остатков.
8. Назовите причины, вызывающие автокорреляцию остатков.
9. Каковы последствия автокорреляции остатков?
10. Опишите методы обнаружения автокорреляции остатков.
11. Опишите методы устранения автокорреляции остатков

***Тема 6. Регрессионные модели с переменной структурой (фиктивные переменные)***

ANOVA-модели. ANCOVA-модели. Использование фиктивных переменных для описания качественных переменных, имеющих более двух альтернатив. Сезонная корректировка модели. Моделирование структурных изменений.

*Вопросы для самоконтроля:*

1. Дайте определение фиктивной переменной.
2. Как записываются значения фиктивной переменной?
3. Какое значение качественной переменной называется «базовым»?
4. Что показывают дифференциальный свободный член и дифференциальный угловой коэффициент?
5. Как производится сезонная корректировка модели с помощью фиктивных переменных?

***Тема 7. Временные ряды***

Характеристики временных рядов. Модели стационарных и нестационарных временных рядов, их идентификация. Моделирование тенденции временного ряда. Моделирование сезонных и циклических колебаний. Оценка моделей с лагами в независимых переменных. Авторегрессионные модели.

*Вопросы для самоконтроля:*

1. Перечислите основные составляющие временного ряда.
2. Перечислите основные виды трендов.
3. Перечислите возможные способы выявления и устранения сезонных колебаний.
4. Каковы основные причины лагов в эконометрических моделях?

***Тема 8. Системы линейных одновременных уравнений***

Общие понятия. Структурная и приведенная формы модели. Проблема идентификации. Оценивание параметров структурной модели. Применение систем уравнений. Косвенный, двухшаговый и трехшаговый метод наименьших квадратов

*Вопросы для самоконтроля:*

1. Каковы основные причины использования систем одновременных уравнений?
2. Назовите возможные способы построения систем уравнений.
3. Как связаны между собой структурная и приведенная формы модели?
4. Раскройте суть косвенного метода наименьших квадратов.
5. В каких случаях используется двухшаговый метод наименьших квадратов?

# 3 Требования, предъявляемые к выполнению и защите курсовойработы

В процессе изучения курса студентам заочной формы обучения предлагается курсовая работа ***на тему*** ***«Построение регрессионной модели экономического объекта»***.

***Целью*** курсовой работы является:

• углубление, закрепление и конкретизация теоретических знаний по данной дисциплине;

• выработка у студентов умения делать обоснованный выбор типа регрессионной модели, методов оценивания параметров, выводы по результатам.

Курсовая работа предполагает ***ответ на теоретический вопрос и выполнение практического задания***, которое включает в себя:

* исследование зависимости между переменными;
* построение первоначальной регрессионной модели экономического объекта;
* проверку качества построенной модели, определение возможностей повышения качества;
* изменение спецификации модели и расчет параметров новой модели;
* проверку качества новой регрессионной модели, сравнение качества моделей;
* оценку прогнозируемого значения моделируемой переменной.

Более подробные задания для каждого варианта курсовой работы изложены в п. 5.

***Вариант задания*** определяется преподавателем индивидуально для каждого студента.

Для выполнения курсовой работы студентам необходимо воспользоваться рекомендуемой в настоящих методических указаниях литературой.

Курсовую работу следует выполнять на листах формата А4 с использованием текстового редактора MS Word и редактора электронных таблиц MS Excel.

Пример оформления титульного листа курсовой работы представлен в Приложении.

Объем курсовой работы и выделение основных разделов определяется студентом самостоятельно, исходя из логики изложения материала.

Для каждого из рассчитываемых в работе показателей следует привести полное наименование и формулу расчета. Если показатель рассчитан с использованием встроенных функций MS Excel, в качестве формулы расчета следует указать применяемую функцию MS Excel с перечислением всех необходимых аргументов.

Курсовая работа должна быть выполнена в сроки, установленные деканатом и преподавателем. Курсовая работа, содержащая ошибки, возвращается студенту на доработку. Курсовая работа, не содержащая ошибок, допускается преподавателем к защите.

***Форма контроля***: защита курсовой работы, включающая ответы на теоретические вопросы, связанные с тематикой курсовой работы, и пояснения по выполнению практических заданий.

# 4 Методические указания по выполнению курсовой работы

Перед выполнением курсовой работы убедитесь, что вы знаете о:

* способах выявления корреляционной взаимосвязи между переменными;
* регрессионных зависимостях между переменными и регрессионных моделях;
* применении метода наименьших квадратов (МНК) для расчета коэффициентов линейного уравнения парной и множественной регрессии;
* показателях качества эконометрических моделей и их интерпретации;
* способах анализа статистической значимости коэффициентов уравнения регрессии и проверки адекватности уравнения регрессии;
* построении доверительных интервалов;
* спецификации регрессионных моделей, ошибках спецификации и методах их обнаружения;
* основных нелинейных регрессионных моделях, и их применении для моделирования экономических объектов;
* точечном и интервальном прогнозировании значений результирующего показателя.

***Внимание!*** Округлять рассчитанные значения показателей необходимо до трех или более десятичных знаков.

При выполнении курсовой работы учитывайте приведенные ниже рекомендации.

## 4.1 Исследование взаимосвязи между переменными

Наглядным способом обнаружить взаимосвязь между двумя переменными является графический способ – построение *поля корреляции*. Каждая точка на этом графике соответствует одному наблюдению – по горизонтали откладывается значение одной из переменных (Х), полученное в результате наблюдения, по вертикали – значение второй переменной (Y), полученное в результате того же наблюдения (см. рисунок 1).

Поле корреляции позволяет приблизительно определить, есть ли корреляционная взаимосвязь между этими переменными, а также вид этой взаимосвязи – линейная или нелинейная, прямая или обратная, сильная или слабая. Однако при использовании поля корреляции невозможно дать количественную оценку силе взаимосвязи.

Сила линейной взаимосвязи характеризуется *коэффициентом корреляции*, который рассчитывается по формуле:

 (1)

где ‾*x* – среднее значение переменной Х (выборочное среднее);

 ‾*y* – среднее значение переменной Y (выборочное среднее);

  - среднее значение произведений ;

 σх и σy – выборочные средние квадратичные отклонения переменных *X* и *Y*:

,  (2)

Коэффициент корреляции можно рассчитать с помощью встроенной функции MS Excel:

*КОРРЕЛ (массив\_Х, массив\_Y),*

где *массив\_Х* – массив значений переменной Х;

 *массив\_Y* – массив значений переменной Y.

Коэффициент корреляции изменяется в диапазоне от -1 до 1. Абсолютное значение коэффициента показывает силу линейной взаимосвязи, а его знак – направление взаимосвязи.

Если значение коэффициента корреляции приблизительно равно 0, то это говорит об отсутствии линейной взаимосвязи между переменными. Однако в этом случае возможно наличие нелинейной взаимосвязи.

Более подробная интерпретация коэффициента корреляции приведена в таблице 2:

Таблица 2 – Интерпретация коэффициента корреляции

|  |  |  |
| --- | --- | --- |
| **Коэффициент корреляции** | **Интерпретация** | **Альтернативная интерпретация** |
| Близко к 1(более 0,9) | Сильная прямая линейная взаимосвязь. Точки на поле корреляции довольно плотно сгруппированы вокруг прямой линии, направленной вверх и вправо. | - |
| 0,6 – 0,9 | Прямая линейная связь средней силы. Точки на поле корреляции сгруппированы вокруг прямой линии, направленной вверх и вправо, но имеют некоторый разброс. | Нелинейная взаимосвязь: точки на поле корреляции сгруппированы вокруг кривой, направленной вверхМежду точками на поле корреляции сильная линейная связь, но один выброс исказил корреляциюМежду точками на поле корреляции в основном связи нет, но один выброс исказил корреляциюКорреляция искажена наличием в составе данных групп по-разному взаимосвязанных объектов |
| 0,3 – 0,6 | Слабая прямая линейная связь. Точки на поле корреляции сгруппированы вокруг прямой линии, направленной вверх и вправо, но имеют значительный разброс. |
| Близко к 0, положительно | Незначительная прямая линейная связь. Точки на поле корреляции образуют случайное облако с незначительной ориентацией вверх и вправо. |
| 0 | Отсутствие линейной взаимосвязи, совершенно случайное облако точек на поле корреляции | Точки на поле корреляции располагаются строго на кривой, поднимающейся с одной стороны и опускающейся с другой стороны |
| Близко к 0, отрицательно | Незначительная обратная линейная связь. Точки на поле корреляции образуют случайное облако с незначительной ориентацией вниз и вправо. | Нелинейная взаимосвязь: точки на поле корреляции сгруппированы вокруг кривой, направленной внизМежду точками на поле корреляции сильная линейная связь, но один выброс исказил корреляциюМежду точками на поле корреляции в основном связи нет, но один выброс исказил корреляциюКорреляция искажена наличием в составе данных групп по-разному взаимосвязанных объектов |
| -0,3 – (-0,6) | Слабая обратная линейная связь. Точки на поле корреляции сгруппированы вокруг прямой линии, направленной вниз и вправо, но имеют значительный разброс. |
| -0,6 – (-0,9) | Обратная линейная связь средней силы. Точки на поле корреляции сгруппированы вокруг прямой линии, направленной вниз и вправо, но имеют некоторый разброс. |
| Близко к -1(менее -0,9) | Сильная обратная линейная взаимосвязь. Точки на поле корреляции плотно сгруппированы вокруг прямой линии, направленной вниз и вправо. | - |

***Пример 1.***

В таблице 3 приведены значения случайных величин *X* (расходы на рекламу, млн. руб.) и *Y* (прибыль предприятия, млн. руб.):

Таблица 3 – Исходные данные

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *X* | 1 | 0,8 | 0,5 | 2,5 | 1,8 | 4 | 5,7 | 9 | 8,3 | 8,8 | 7,5 |
| *Y* | 7 | 5 | 3 | 15 | 13 | 20 | 25 | 15 | 20 | 17 | 22 |

Наличие и степень линейной зависимости между переменными У и Х оценим с помощью коэффициента корреляции, используя формулы (1) и (2). Промежуточные вычисления представлены в таблице 4.

Таблица 4 – Промежуточные вычисления для расчета коэффициента корреляции

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***i*** | ***xi*** | ***yi*** | ***xi2*** | ***xiyi*** | ***yi2*** |
| 1 | 1,0 | 7 | 1,00 | 7,0 | 49 |
| 2 | 0,8 | 5 | 0,64 | 4,0 | 25 |
| 3 | 0,5 | 3 | 0,25 | 1,5 | 9 |
| 4 | 2,5 | 15 | 6,25 | 37,5 | 225 |
| 5 | 1,8 | 13 | 3,24 | 23,4 | 169 |
| 6 | 4,0 | 20 | 16,00 | 80,0 | 400 |
| 7 | 5,7 | 25 | 32,49 | 142,5 | 625 |
| 8 | 9,0 | 15 | 81,00 | 135,0 | 225 |
| 9 | 8,3 | 20 | 68,89 | 166,0 | 400 |
| 10 | 8,8 | 17 | 77,44 | 149,6 | 289 |
| 11 | 7,5 | 22 | 56,25 | 165,0 | 484 |
| *Сумма* | 49,9 | 162 | 343,45 | 911,5 | 2900 |
| *Среднее* | 4,536 | 14,727 | 31,223 | 82,864 | 263,637 |

Используя рассчитанные в таблице 4 значения средних величин, получим:







Данное значение коэффициента корреляции свидетельствует о *прямой линейной связи средней силы* между переменными *X* и *Y*, т.е. точки на поле корреляции сгруппированы вокруг прямой линии, направленной вверх и вправо, но имеют некоторый разброс (см. таблицу 2).

Альтернативная интерпретация коэффициента корреляции – нелинейная взаимосвязь между переменными, либо наличие выбросов – отдельных точек, отклоняющихся от общей совокупности.

Построим поле корреляции (см. рисунок 1).

По расположению точек на поле корреляции предполагаем, что между переменными *X* и *Y* наблюдается нелинейная корреляционная зависимость.

Можно предположить, что зависимость между переменными имеет характер параболы.

Для описания зависимости между переменными, кроме параболы, можно предлагать и другие математические функции – гиперболу, логарифмическую, степенную или экспоненциальную функцию. Предлагаемая функция должна наилучшим образом соответствовать расположению точек на поле корреляции.

*Рисунок 1 – Поле корреляции*

## 4.2 Построение уравнения парной линейной регрессии

Линейная регрессия находит широкое применение в эконометрике ввиду четкой экономической интерпретации ее параметров.

Уравнение парной линейной регрессии позволяет предсказывать значения *результирующего показателя* (переменной Y) на основании значений единственного *фактора* (переменной Х), который связан с результирующим показателем линейной зависимостью:

Ŷ = b0 + b1 · X (3)

Построение уравнения парной линейной регрессии сводится к оценке его *параметров* – углового коэффициента b1 и свободного члена b0.

Оценки параметров линейной регрессии могут быть найдены разными способами. Классический подход к оцениванию параметров линейной регрессии основан на методе наименьших квадратов (МНК). МНК позволяет получить такие оценки параметров b1 и b0, при которых сумма квадратов отклонений фактических значений результирующего показателя yi от расчетных (теоретических) значений ŷi минимальна. Для этого используются следующие формулы:

 (4)

Угловой коэффициент b1 можно рассчитать с помощью встроенной функции MS Excel:

*НАКЛОН (известные значения\_Y, известные значения\_Х),*

где *известные значения\_Y* – массив значений переменной Y;

 *известные значения\_Х* – массив значений переменной Х.

Свободный член b0 можно рассчитать с помощью встроенной функции MS Excel:

*ОТРЕЗОК (известные значения\_Y, известные значения\_Х),*

где *известные значения\_Y* – массив значений переменной Y;

 *известные значения\_Х* – массив значений переменной Х.

Кроме того, коэффициенты уравнения регрессии b0 и b1 входят в состав показателей, рассчитываемых функцией ЛИНЕЙН (см. Приложение В).

После определения параметров (коэффициентов) уравнения регрессии, необходимо записать построенное уравнение в виде (3) и дать интерпретацию его параметров:

- *угловой коэффициент* b1 показывает изменение результирующего показателя Y при изменении фактора Х на единицу;

*- свободный член* b0 определяет значение результирующего показателя Y при нулевом значении фактора Х.

***Пример 2.***

Построим уравнение парной линейной регрессии для описания зависимости прибыли предприятия от расходов на рекламу, используя исходные данные из таблицы 3 (см. пример 1).

В данном случае прибыль предприятия рассматривается как *результирующий показатель* (Y), а расходы на рекламу как *фактор* (Х).

Для определения параметров (коэффициентов) уравнения регрессии используем формулы метода наименьших квадратов (4) и рассчитанные в таблице 4 значения средних величин:



Таким образом, уравнение парной линейной регрессии имеет вид:

*Ŷ = 7,885 + 1,508⋅Х*

Экономическая интерпретация параметров уравнения регрессии:

а) угловой коэффициент *b1* показывает, что при возрастании расходов на рекламу на 1 млн. руб. прибыль предприятия возрастет на 1,508 млн. руб.;

б) свободный член *b0* определяет прогнозируемое значение прибыли предприятия при нулевых расходах на рекламу: 7,885 млн. руб.

## 4.3 Проверка качества регрессионной модели

После построения уравнения регрессии необходимо проверить, является ли оно качественным, так как только в этом случае уравнение можно использовать для прогнозирования и анализа взаимосвязи между экономическими переменными.

Проверка качества регрессионной модели должна содержать:

* проверку адекватности уравнения регрессии эмпирическим данным;
* проверку адекватности уравнения регрессии данным генеральной совокупности;
* проверку статистической значимости коэффициентов уравнения регрессии.

Кроме того, необходимо проверить, соответствуют ли свойства случайного отклонения определенным требованиям (подробнее см. п. 4.4).

Каждый этап проверки качества должен сопровождаться подробными выводами.

## 4.3.1 Проверка адекватности уравнения регрессии эмпирическим данным

*Адекватностью* называется соответствие уравнения реальным наблюдениям.

*Эмпирическими* называются данные наблюдений, используемые при построении уравнения (данные выборки). Таким образом, проверка адекватности уравнения регрессии эмпирическим данным заключается в определении степени соответствия уравнения данным выборки.

В случае парной регрессии можно графически оценить адекватность уравнения эмпирическим данным как соответствие линии регрессии точкам на поле корреляции. Однако точные выводы можно сделать только с помощью *аналитических показателей*:

1) ***стандартная ошибка регрессии***, определяемая по формуле:

, (5)

где *n* – количество наблюдений (объем выборки);

 *ei* – *случайное отклонение* результирующего показателя:

еi = yi - ŷi (6)

Стандартная ошибка регрессии показывает, на сколько в среднем отклоняются наблюдаемые значения результирующего показателя от уравнения регрессии. Чем меньше величина стандартной ошибки, тем более адекватным является уравнение регрессии.

2) ***коэффициент детерминации***, рассчитываемый по формуле:

 (7)

Коэффициент детерминации показывает, какую долю вариации (разброса) результирующего показателя Y можно объяснить с помощью уравнения регрессии (т.е. действием фактора Х).

Коэффициент детерминации может принимать значения от 0 до 1. Чем ближе коэффициент детерминации к 1, тем большая доля вариации результирующего показателя объясняется действием фактора Х, т.е. тем более адекватным является уравнение регрессии.

В Приложении Б приведены функции MS Excel, помогающие рассчитать приведенные выше показатели (функции вычисления суммы значений, среднего значения и квадратного корня).

Кроме того, стандартная ошибка регрессии Se и коэффициент детерминации R2 входят в состав показателей, рассчитываемых функцией ЛИНЕЙН (см. Приложение В).

***Пример 3.***

Проверим качество уравнения парной линейной регрессии, построенного в примере 2:

*Ŷ = 7,885 + 1,508⋅Х*

Используя построенное уравнение, рассчитаем прогнозируемые значения результирующего показателя *ŷi*. Для этого необходимо подставить в уравнение каждое из значений фактора *xi*. Запишем полученные значения в таблицу 5.

Построим линию регрессии на поле корреляции, используя рассчитанные теоретические значения результирующего показателя *ŷi* (см. рисунок 2).

На рисунке видно, что линия регрессии недостаточно хорошо соответствует полю корреляции. Однако линия все же отражает общую тенденцию к увеличению переменной Y.

*Рисунок 2 – Линия регрессии*

Таблица 5 – Промежуточные расчеты для проверки адекватности уравнения регрессии

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***i*** | ***xi*** | ***yi*** | ***ŷi*** | ***ei*** | ***ei2*** |
| 1 | 1,0 | 7 | 9,393 | -2,393 | 5,727 |
| 2 | 0,8 | 5 | 9,091 | -4,091 | 16,740 |
| 3 | 0,5 | 3 | 8,639 | -5,639 | 31,797 |
| 4 | 2,5 | 15 | 11,656 | 3,344 | 11,185 |
| 5 | 1,8 | 13 | 10,600 | 2,400 | 5,761 |
| 6 | 4,0 | 20 | 13,918 | 6,082 | 36,988 |
| 7 | 5,7 | 25 | 16,482 | 8,518 | 72,548 |
| 8 | 9,0 | 15 | 21,460 | -6,460 | 41,733 |
| 9 | 8,3 | 20 | 20,404 | -0,404 | 0,163 |
| 10 | 8,8 | 17 | 21,158 | -4,158 | 17,293 |
| 11 | 7,5 | 22 | 19,198 | 2,802 | 7,854 |
| *Сумма* | - | 162 | - | - | 247,789 |
| *Среднее* | - | 14,7273 | - | - | - |

Проверим адекватность уравнения регрессии эмпирическим данным на основе аналитических показателей – стандартной ошибки регрессии и коэффициента детерминации.

Результаты промежуточных расчетов запишем в таблицу 5.

Рассчитаем стандартную ошибку регрессии:



Стандартная ошибка регрессии показывает, что наблюдаемые значения прибыли предприятия yi отклоняются от прогнозов, рассчитанных с помощью уравнения регрессии, в среднем на 5,247 млн. руб. Таким образом, ошибка достаточно велика.

Рассчитаем коэффициент детерминации, используя вторую часть формулы (7):



Значение коэффициента детерминации R2 = 0,518 показывает, что 51,8% разброса значений прибыли предприятия объясняется с помощью уравнения регрессии (изменением расходов на рекламу), а 48,2% объясняется случайными причинами.

Таким образом, построенное уравнение объясняет около половины разброса значений результирующего показателя, следовательно, *может быть признано адекватным эмпирическим данным*. Хотя адекватность данного уравнения достаточно низкая, так как коэффициент детерминации намного меньше 1.

Вывод об адекватности уравнения на основании значения коэффициента детерминации зависит от количества наблюдений: например, при R2 = 0,2 и количестве наблюдений n ≤ 20 уравнение нельзя признать адекватным, а при большем количестве наблюдений такое значение коэффициента может говорить об адекватности уравнения.

Можно воспользоваться следующими рекомендациями:

- для выборки объема n ≈ 10 уравнение следует признать адекватным эмпирическим данным при R2 > 0,4;

- для выборки объема n ≈ 20 уравнение следует признать адекватным эмпирическим данным при R2 > 0,2;

- для выборки объема n ≈ 30 уравнение следует признать адекватным эмпирическим данным при R2 > 0,15.

## 4.3.2 Проверка адекватности уравнения регрессии данным генеральной совокупности

Различные произвольные значения, которые могут принимать переменные Х и Y, называются *теоретическими данными* (данными генеральной совокупности).

Уравнение линейной регрессии является адекватным данным генеральной совокупности, если существует значимая линейная взаимосвязь между фактором и результирующим показателем. Таким образом, проверка адекватности линейного уравнения сводится к проверке значимости линейной взаимосвязи между переменными.

Для этого необходимо проверить ***статистическую гипотезу о значимости коэффициента детерминации*** ***R2***.

Нулевая и альтернативная гипотезы формулируются следующим образом:

Н0: R2 = 0 – значение коэффициента детерминации в генеральной совокупности незначимо, связь между фактором и результирующим показателем отсутствует;

Н1: R2 > 0 – значение коэффициента детерминации в генеральной совокупности значимо, существует значимая связь между фактором и результирующим показателем.

Для проверки используется *критерий Фишера*:

, (8)

где *n* – количество наблюдений (объем выборки).

Критерий F имеет распределение Фишера с *числами степеней свободы* ν1 = 1 и ν2 = n-2.

По таблице критических точек Фишера [1,3,6] определяется значение *критической точки fкр*, которое зависит от следующих показателей:

- уровня значимости α, значение которого задается при проверке гипотезы;

- числа степеней свободы ν1;

- числа степеней свободы ν2.

Чтобы сделать вывод о значимости коэффициента детерминации, необходимо сравнить между собой рассчитанное значение критерия F и значение критической точки fкр.

Если значение критерия F попадает в *область принятия гипотезы* (F < fкр), можно сделать следующие выводы:

1. коэффициент детерминации в генеральной совокупности признается незначимым;
2. в генеральной совокупности отсутствует значимая линейная связь между фактором и результирующим показателем;
3. построенное уравнение неадекватно данным генеральной совокупности.

Если значение критерия F попадает в *критическую область* (F ≥ fкр), выводы оказываются следующими:

1. коэффициент детерминации в генеральной совокупности признается значимым;
2. в генеральной совокупности существует значимая линейная связь между фактором и результирующим показателем;
3. построенное уравнение адекватно данным генеральной совокупности.

Критическую точку fкр можно определить с помощью встроенной функции MS Excel:

*F.ОБР.ПХ (вероятность; степени\_свободы1, степени\_свободы2),*

где *вероятность* – уровень значимости α;

 *степени\_свободы1* – число степеней свободы ν1;

 *степени\_свободы2* – число степеней свободы ν2.

Критерий Фишера F входит в состав показателей, рассчитываемых функцией ЛИНЕЙН (см. Приложение В).

***Пример 4.***

Проверим адекватность уравнения *Ŷ = 7,885 + 1,508⋅Х*, построенного в примере 2, данным генеральной совокупности.

Для этого проверим значимость коэффициента детерминации в генеральной совокупности при уровне значимости α = 0,05, используя значение коэффициента детерминации R2 = 0,518, рассчитанное в примере 3.

Сформулируем нулевую и альтернативную гипотезы:

Н0: R2 = 0 – значение коэффициента детерминации в генеральной совокупности незначимо, связь между фактором и результирующим показателем отсутствует;

Н1: R2 > 0 – значение коэффициента детерминации в генеральной совокупности значимо, существует значимая связь между фактором и результирующим показателем.

Рассчитаем критерий Фишера по формуле (8):



Критерий F имеет числа степеней свободы ν1 = 1 и ν2 = n – 2 = 11 – 2 = 9.

Определим критическую точку fкр по таблице критических точек Фишера (таблицы критических точек можно найти в учебной литературе [1,3,6] и справочных изданиях по эконометрике, а также в раздаточном материале к курсовой работе на сайте moodle.nirhtu.ru):

при уровне значимости α = 0,05 и числах степеней свободы ν1 = 1 и ν2 = 9 критическая точка fкр = 5,117.

Поскольку F > fкр (9,676 > 5,117), можно сделать вывод о том, что коэффициент детерминации R2 в генеральной совокупности является значимым. Таким образом, в генеральной совокупности существует значимая линейная связь между прибылью предприятия и расходами на рекламу, и построенное уравнение регрессии адекватно данным генеральной совокупности.

## 4.3.3 Проверка статистической значимости коэффициентов регрессии

Уравнение регрессии Ŷ = b0 + b1 · X, построенное по эмпирическим данным (данным выборки), называется *эмпирическим уравнением регрессии*.

Если бы можно было изучить все возможные значения переменных (данные генеральной совокупности), можно было бы построить *теоретическое уравнение регрессии*:

M(Y|X) = β0 + β1 · X, (9)

где M(Y|X) – условное математическое ожидание переменной Y;

 β0 – теоретический свободный член регрессии;

 β1 – теоретический угловой коэффициент регрессии.

На практике построить теоретическое уравнение регрессии невозможно.

Эмпирическое уравнение регрессии является *оценкой* теоретического уравнения, т.е. отражает его основные свойства, но имеет некоторые отличия. В том числе и эмпирические коэффициенты регрессии b0 и b1 могут отличаться от теоретических коэффициентов β0 и β1.

Поэтому проверку качества уравнения регрессии необходимо дополнить анализом возможных значений теоретических коэффициентов, одним из направлений которого является проверка статистической значимости теоретических коэффициентов регрессии.

Проверка статистической значимости коэффициентов производится с помощью аппарата статистической проверки гипотез, и позволяет определить с заданной надежностью, могут ли теоретические коэффициенты принимать нулевые значения.

Наиболее важной является проверка статистической значимости углового коэффициента, поскольку он отражает влияние фактора Х на результирующий показатель Y. Если угловой коэффициент незначим, уравнение регрессии преобразуется в уравнение вида Ŷ = β0, т.е. значения результирующего показателя не зависят от значений фактора, а следовательно, линейное уравнение регрессии нельзя использовать для прогнозирования в генеральной совокупности.

Если угловой коэффициент значим, это говорит о наличии линейной связи между переменными Х и Y в генеральной совокупности, т.е. подтверждает вывод об адекватности уравнения регрессии данным генеральной совокупности.

***Статистическая гипотеза о значимости теоретического углового коэффициента β1:***

Нулевая и альтернативная гипотезы формулируются следующим образом:

Н0: β1 = 0 – значение теоретического углового коэффициента незначимо, связь между фактором и результирующим показателем отсутствует;

Н1: β1 ≠ 0 – значение теоретического углового коэффициента значимо, существует значимая связь между фактором и результирующим показателем.

Для проверки используется критерий Стьюдента:

, (10)

где b1 – эмпирический угловой коэффициент регрессии;

Sb1 – стандартная ошибка углового коэффициента регрессии, которая определяется по формуле:

, (11)

где Sе – стандартная ошибка регрессии (см. п. 4.3.1);

 *n* – количество наблюдений (объем выборки).

Критерий tb1 имеет распределение Стьюдента с числом степеней свободы ν = n – 2.

По таблице критических точек Стьюдента [1,3,6] определяется значение *критической точки tкр*, которое зависит от следующих показателей:

- уровня значимости α, значение которого задается при проверке гипотезы;

- числа степеней свободы ν.

Чтобы сделать вывод о значимости теоретического углового коэффициента регрессии β1, необходимо сравнить между собой рассчитанное значение критерия tb1 и значение критической точки tкр.

Если значение критерия tb1 попадает в *область принятия гипотезы* (-tкр < tb1 < tкр), можно сделать следующие выводы:

1. теоретический угловой коэффициент признается незначимым;
2. в генеральной совокупности отсутствует значимая линейная связь между фактором и результирующим показателем;
3. построенное уравнение неадекватно данным генеральной совокупности.

Если значение критерия tb1 попадает в *одну из критических областей* (tb1 ≤ -tкр или
tb1 ≥ tкр), выводы оказываются следующими:

1. теоретический угловой коэффициент признается значимым;
2. в генеральной совокупности существует значимая линейная связь между фактором и результирующим показателем;
3. построенное уравнение адекватно данным генеральной совокупности.

***Статистическая гипотеза о значимости теоретического свободного члена β0:***

Нулевая и альтернативная гипотезы формулируются следующим образом:

Н0: β0 = 0 – значение теоретического свободного члена незначимо;

Н1: β0 ≠ 0 – значение теоретического свободного члена значимо.

Для проверки используется критерий Стьюдента:

, (12)

где b0 – эмпирический свободный член регрессии;

Sb0 – стандартная ошибка свободного члена регрессии, которая определяется по формуле:

, (13)

где Sb1 – стандартная ошибка углового коэффициента регрессии, рассчитанная по формуле (11).

Критерий tb0 имеет распределение Стьюдента с числом степеней свободы ν = n – 2.

Значение *критической точки tкр* будет таким же, как и при проверке значимости углового коэффициента. Оно определяется по таблице критических точек Стьюдента [1,3,6] в зависимости от следующих показателей:

- уровня значимости α, значение которого задается при проверке гипотезы;

- числа степеней свободы ν.

Чтобы сделать вывод о значимости теоретического свободного члена регрессии β0, необходимо сравнить между собой рассчитанное значение критерия tb0 и значение критической точки tкр.

Если значение критерия tb0 попадает в *область принятия гипотезы* (-tкр < tb0 < tкр), можно сделать следующие выводы:

1. теоретический свободный член признается незначимым;
2. линия регрессии может проходить через начало координат.

Если значение критерия tb0 попадает в *одну из критических областей* (tb0 ≤ -tкр или
tb0 ≥ tкр), выводы оказываются следующими:

1. теоретический свободный член признается значимым;
2. линия регрессии не может проходить через начало координат.

Критическую точку tкр можно определить с помощью встроенной функции MS Excel:

*СТЬЮДЕНТ.ОБР.2Х (вероятность; степени\_свободы),*

где *вероятность* – уровень значимости α;

 *степени\_свободы* – число степеней свободы ν.

Стандартные ошибки коэффициентов регрессии Sb1 и Sb0 входят в состав показателей, рассчитываемых функцией ЛИНЕЙН (см. Приложение В).

***Пример 5.***

Проверим значимость теоретических коэффициентов линейного уравнения, используя эмпирическое уравнение *Ŷ = 7,885 + 1,508⋅Х*, построенное в примере 2.

Проверим значимость углового коэффициента регрессии β1 при уровне значимости α = 0,05.

Сформулируем нулевую и альтернативную гипотезы:

Н0: β1 = 0 – значение теоретического углового коэффициента незначимо, связь между фактором и результирующим показателем отсутствует;

Н1: β1 ≠ 0 – значение теоретического углового коэффициента значимо, существует значимая связь между фактором и результирующим показателем.

Рассчитаем стандартную ошибку углового коэффициента, используя вторую часть формулы (11). Значение стандартной ошибки регрессии рассчитано в примере 3: Se = 5,247. Средние значения для переменной Х рассчитаны в таблице 4. Тогда:



Рассчитаем критерий Стьюдента по формуле (10):



Критерий tb1 имеет число степеней свободы ν = n – 2 = 11 – 2 = 9.

Определим критическую точку tкр по таблице критических точек Стьюдента (таблицы критических точек можно найти в учебной литературе [1,3,6] и справочных изданиях по эконометрике, а также в раздаточном материале к курсовой работе на сайте moodle.nirhtu.ru):

при уровне значимости α = 0,05 и числе степеней свободы ν = 9 критическая точка tкр = 2,262.

Поскольку tb1 > tкр (3,111 > 2,262), можно сделать вывод о том, что теоретический угловой коэффициент регрессии является значимым. Таким образом, в генеральной совокупности существует значимая линейная связь между прибылью предприятия и расходами на рекламу, и построенное уравнение регрессии адекватно данным генеральной совокупности.

Проверим значимость свободного члена регрессии β0 при уровне значимости α = 0,05.

Сформулируем нулевую и альтернативную гипотезы:

Н0: β0 = 0 – значение теоретического свободного члена незначимо;

Н1: β0 ≠ 0 – значение теоретического свободного члена значимо.

Рассчитаем стандартную ошибку свободного члена по формуле (13). Средний квадрат значений переменной Х рассчитан в таблице 4: ‾х2 = 31,223. Тогда:



Рассчитаем критерий Стьюдента по формуле (12):



Критерий tb0 имеет число степеней свободы ν = n – 2 = 11 – 2 = 9.

Критическая точка tкр имеет то же значение, что и при проверке значимости углового коэффициента:

при уровне значимости α = 0,05 и числе степеней свободы ν = 9 критическая точка tкр = 2,262.

Поскольку tb0 > tкр (2,910 > 2,262), можно сделать вывод о том, что теоретический свободный член регрессии является значимым. Следовательно, линия регрессии не может проходить через начало координат.

## 4.4 Проверка свойств случайного отклонения

При построении уравнения регрессии предполагается, что выполняются определенные условия относительно свойств случайного отклонения εi в генеральной совокупности, среди которых:

- условие постоянства дисперсии εi;

- условие некоррелированности значений εi и др.

Совокупность этих условий называется *предпосылками метода наименьших квадратов* (МНК)или *условиями Гаусса-Маркова* [1, 2, 4, 6].

Уравнение регрессии считается качественным только в том случае, если все предпосылки МНК действительно выполняются. Поэтому после построения уравнения регрессии необходимо выполнить проверку свойств случайного отклонения по следующим направлениям:

- проверка наличия корреляции между значениями εi (автокорреляции остатков);

- проверка постоянства дисперсии εi (проверка наличия гомоскедастичности или гетероскедастичности).

## 4.4.1 Автокорреляция остатков

*Автокорреляцией остатков* называется зависимость между значениями случайных отклонений εi, упорядоченными по значениям фактора *Х*.

Может наблюдаться автокорреляция остатков двух видов:

- *положительная* автокорреляция возникает, если большинство соседних случайных отклонений имеют одинаковый знак (оба положительны или оба отрицательны);

- *отрицательная* автокорреляция возникает, если большинство соседних случайных отклонений имеют различные знаки.

Наличие автокорреляции остатков является нарушением одной из предпосылок МНК, поэтому при обнаружении автокорреляции необходимо выявить ее причину и устранить ее.

Существует несколько *методов обнаружения автокорреляции*:

- графический метод;

- метод рядов;

- тест Дарбина-Уотсона.

***1 Графический метод обнаружения автокорреляции остатков***

Этот метод заключается в построении *диагностической диаграммы* – точечной графической зависимости между прогнозируемыми значениями результирующего показателя ŷi и эмпирическими случайными отклонениями ei (в случае парной регрессии можно также рассматривать зависимость случайных отклонений ei от значений фактора xi).

В случае положительной автокорреляции на диагностической диаграмме видна явная зависимость между точками, обычно нелинейная (см. рисунок 3 а,б).

а) б)

в)

*Рисунок 3 – Примеры диагностических диаграмм при наличии автокорреляции остатков*

В случае отрицательной автокорреляции можно обнаружить, что случайные отклонения слишком часто меняют знак, т.е. присутствуют частые колебания относительно оси абсцисс см. рисунок 3 в).

Если явная взаимосвязь между случайными отклонениями отсутствует (точки на диагностической диаграмме при этом образуют бесформенное облако), то необходимо признать отсутствие автокорреляции остатков.

***Пример 6.***

Проверим графическим методом наличие автокорреляции остатков в линейной регрессионной модели, построенной в примере 2.

Для этого построим диагностическую диаграмму (см. рисунок 4) по данным таблицы 5: на оси абсцисс отложим прогнозируемые значения результирующего показателя ŷi, а на оси ординат – соответствующие им значения случайного отклонения ei.

*Рисунок 4 – Диагностическая диаграмма*

На построенной диагностической диаграмме заметна явная нелинейная взаимосвязь между точками, причем кривая случайных отклонений пересекает ось всего два раза, значит, наблюдается положительная автокорреляция остатков в регрессионной модели.

***2 Метод рядов для обнаружения автокорреляции остатков***

Этот метод заключается в выполнении следующей последовательности шагов:

1) наблюдения упорядочиваются по возрастанию значений фактора xi (при множественной регрессии – по возрастанию прогнозируемых значений результирующего показателя ŷi);

2) рассматривается последовательность знаков эмпирических случайных отклонений еi и определяется количество рядов *k* (ряд – последовательность случайных отклонений с одинаковым знаком);

3) определяется общее количество положительных случайных отклонений *n1* и общее количество отрицательных случайных отклонений *n2*;

4) по таблице критических точек для метода рядов [1] при заданном уровне значимости *α* в зависимости от *n1* и *n2* определяется *k1* – нижняя граница количества рядов и *k2* – верхняя граница количества рядов;

5) делается вывод:

- при *k  k1* наблюдается положительная автокорреляция остатков (рядов слишком мало по сравнению с количеством наблюдений, т.е. кривая случайных отклонений редко пересекает ось);

- при *k  k2* наблюдается отрицательная автокорреляция остатков (рядов слишком много по сравнению с количеством наблюдений, т.е. кривая случайных отклонений слишком часто пересекает ось);

- при *k1 < k < k2* автокорреляция остатков отсутствует.

***Пример 7.***

Проверим методом рядов при уровне значимости 0,05 наличие автокорреляции остатков в линейной регрессионной модели, построенной в примере 2.

Рассчитанные значения случайных отклонений представлены в таблице 5.

1) Упорядочим значения случайного отклонения по возрастанию значений фактора xi:

Таблица 6 – Упорядочивание значений случайных отклонений

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| xi | 0,5 | 0,8 | 1,0 | 1,8 | 2,5 | 4,0 | 5,7 | 7,5 | 8,3 | 8,8 | 9,0 |
| еi | -5,639 | -4,091 | -2,393 | 2,400 | 3,344 | 6,082 | 8,518 | 2,802 | -0,404 | -4,158 | -6,460 |

2) Рассмотрим последовательность знаков случайных отклонений ei и определим количество рядов:

1 ряд содержит отрицательные случайные отклонения: -5,639; -4,091; -2,393.

2 ряд содержит положительные случайные отклонения: 2,400; 3,344; 6,082; 8,518; 2,802.

3 ряд содержит отрицательные случайные отклонения: -0,404; -4,158; -6,460.

Количество рядов *k=3.*

3) Количество положительных случайных отклонений *n1=5,* количество отрицательных случайных отклонений *n2=6.*

4) Определим нижнюю и верхнюю границы количества рядов *k1* и *k2* по таблице критических точек для метода рядов (таблицы критических точек можно найти в учебной литературе [1], а также в раздаточном материале к курсовой работе на сайте moodle.nirhtu.ru):

при уровне значимости *α = 0,05, n1 = 5* и *n2 = 6* определяем *k1 = 3* и *k2* = *10*.

5) Так как *k = k1*, то делаем вывод о наличии положительной автокорреляции остатков в регрессионной модели.

***3 Тест Дарбина-Уотсона для обнаружения автокорреляции остатков***

Тест основан на вычислении *DW*-критерия по формуле (14):

 (14)

где *еi* – эмпирические случайные отклонения.

Перед расчетом критерия наблюдения необходимо упорядочить по возрастанию значений фактора xi (при множественной регрессии – по возрастанию прогнозируемых значений результирующего показателя ŷi).

Вычисленное значение *DW*-критерия сравнивается с критическими точками. Множество возможных значений *DW*-критерия можно представить в виде пяти интервалов на числовой оси (см. рисунок 5):



*Рисунок 5 – Графическая интерпретация возможных значений критерия DW*

Критические точки *dl* и *du* выбираются из таблицы Дарбина-Уотсона [1,3,6] при заданном уровне значимости *α* в зависимости от следующих показателей:

- числа наблюдений (объема выборки) *n*;

- числа факторов в регрессионной модели *m* (в парной регрессионной модели *m = 1*).

Вывод о наличии автокорреляции делается в зависимости от принадлежности критерия DW одному из интервалов:

- при *0* ≤ *DW* ≤ *dl* наблюдается положительная автокорреляция остатков;

- при *du* ≤ *DW* ≤(*4 – du*) автокорреляция остатков отсутствует;

- при (*4 – dl*) ≤ *DW* ≤ *4* наблюдается отрицательная автокорреляция остатков;

- оставшиеся интервалы (при *dl <DW < du* или(*4 – du*) *< DW <* (*4 – dl*)) являются областями неопределенности – однозначно сделать вывод о наличии или отсутствии автокорреляции остатков в этом случае невозможно.

***Пример 8.***

Проверим с помощью теста Дарбина-Уотсона при уровне значимости 0,05 наличие автокорреляции остатков в линейной регрессионной модели, построенной в примере 2.

Значения случайных отклонений, упорядоченные по возрастанию значений фактора xi, представлены в таблице 6. Выполним промежуточные вычисления (см. таблицу 7) и рассчитаем *DW*-критерий по формуле (14).

Таблица 7 – Промежуточные вычисления для расчета критерия DW

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *i* | *xi* | *ei* | *ei2* | *ei-1* | *ei- ei-1* | *(ei - ei-1)2* |
| 1 | 0,5 | -5,639 | 31,797 | - | - | - |
| 2 | 0,8 | -4,091 | 16,740 | -5,639 | 1,547 | 2,395 |
| 3 | 1 | -2,393 | 5,727 | -4,091 | 1,698 | 2,884 |
| 4 | 1,8 | 2,400 | 5,761 | -2,393 | 4,793 | 22,976 |
| 5 | 2,5 | 3,344 | 11,185 | 2,400 | 0,944 | 0,891 |
| 6 | 4 | 6,082 | 36,988 | 3,344 | 2,737 | 7,494 |
| 7 | 5,7 | 8,518 | 72,548 | 6,082 | 2,436 | 5,933 |
| 8 | 7,5 | 2,802 | 7,854 | 8,518 | -5,715 | 32,662 |
| 9 | 8,3 | -0,404 | 0,163 | 2,802 | -3,207 | 10,283 |
| 10 | 8,8 | -4,158 | 17,293 | -0,404 | -3,754 | 14,094 |
| 11 | 9 | -6,460 | 41,733 | -4,158 | -2,302 | 5,298 |
| *Сумма* | - | - | 247,789 | - | - | 104,909 |

**

Определим критические точки *dl* и *du* по таблице критических точек Дарбина-Уотсона (таблицы критических точек можно найти в учебной литературе [1,3,6], а также в раздаточном материале к курсовой работе на сайте moodle.nirhtu.ru):

при уровне значимости *α = 0,05,* числе наблюдений *n = 11* ичисле факторов *m = 1* определяем *dl = 0,927* и *du* = *1,324*.

Рассчитаем другие критические точки:

4 – du = 4 – 1,324 = 2,676;

4 – dl = 4 – 0,927 = 3,073.

Нанесем критические точки на отрезок возможных значений критерия DW (см. рисунок 6):



*Рисунок 6 – Графическая интерпретация теста Дарбина-Уотсона*

Рассчитанное значение критерия DW попало в интервал *0* ≤ *DW* ≤ *dl* (*0* ≤ *0,423* ≤ *0,927*), следовательно, можно сделать вывод о наличии положительной автокорреляции остатков в регрессионной модели.

При обнаружении автокорреляции остатков необходимо выяснить причины ее возникновения и предложить способ устранения.

Основными *причинами наличия автокорреляции остатков* в регрессионных моделях являются:

- ошибки спецификации: пропуск важной объясняющей переменной, использование ошибочной функциональной зависимости между переменными; для устранения автокорреляции в этом случае необходимо изменить спецификацию модели (подробнее см. п. 4.5);

- ошибки измерений; в этом случае необходимо построить новую регрессионную модель, используя новую выборку;

- характер наблюдений (цикличность, наблюдаемая в поведении многих экономических показателей, сглаживание данных – усреднение данных за некоторый длительный временной период); в этом случае необходимо использовать специальные методы построения эконометрической модели временного ряда.

## 4.4.2 Гомо- и гетероскедастичность

Одной из предпосылок МНК является условие постоянства дисперсии случайных отклонений – *гомоскедастичность*. Нарушение этого условия, т.е. изменение дисперсии случайных отклонений при изменении значений фактора xi называют *гетероскедастичностью*.

Поскольку наличие гетероскедастичности является нарушением одной из предпосылок МНК, при ее обнаружении необходимо выявить ее причину и устранить ее.

Существует несколько *методов обнаружения гетероскедастичности*:

- графический метод;

- тест ранговой корреляции Спирмена;

- тест Голдфельда-Квандта.

***1 Графический метод обнаружения гетероскедастичности***

Этот метод заключается в построении *диагностической диаграммы* – точечной графической зависимости между прогнозируемыми значениями результирующего показателя ŷi и эмпирическими случайными отклонениями ei (в случае парной регрессии можно также рассматривать зависимость случайных отклонений ei от значений фактора xi).

Если дисперсии случайных отклонений приблизительно одинаковые для разных значений *ŷi* (или *xi*), т.е. точки на диаграмме представляют собой полосу вокруг оси абсцисс приблизительно равной ширины, значит, присутствует гомоскедастичность.

Если дисперсии случайных отклонений различаются для разных значений *ŷi* (или *xi*), т.е. разброс точек вокруг оси абсцисс на диаграмме постепенно увеличивается или уменьшается, значит, присутствует гетероскедастичность.

Пример диагностической диаграммы при наличии гетероскедастичности представлен на рисунке 7.

*Рисунок 7 – Пример диагностической диаграммы при наличии гетероскедастичности*

***Пример 9.***

Проверим графическим методом наличие гетероскедастичности в линейной регрессионной модели, построенной в примере 2.

Для этого построим диагностическую диаграмму (см. рисунок 8) по данным таблицы 5: на оси абсцисс отложим прогнозируемые значения результирующего показателя ŷi, а на оси ординат – соответствующие им значения случайного отклонения ei (данная диаграмма повторяет диагностическую диаграмму, построенную при обнаружении автокорреляции остатков, см. рисунок 4; при выполнении курсовой работы повторную диаграмму можно не строить).

*Рисунок 8 – Диагностическая диаграмма*

На диагностической диаграмме заметно небольшое увеличение дисперсии (разброса точек) при увеличении прогнозируемых значений результирующего показателя: при ŷ < 15 значения случайного отклонения находятся в интервале (-6 ; 6), а при ŷ > 15 интервал значений случайного отклонения расширяется до (-9 ; 9). На рисунке 7 границы указанных интервалов показаны штриховыми линиями.

Такое увеличение дисперсии не является существенным, значит, можно сделать вывод о постоянстве дисперсии случайных отклонений, т.е. о наличии гомоскедастичности в построенной регрессионной модели (или об отсутствии гетероскедастичности).

***2 Тест ранговой корреляции Спирмена для обнаружения гетероскедастичности***

Суть теста ранговой корреляции Спирмена сводится к оценке коэффициента корреляции между рангами значений переменной Х (для множественной регрессии Ŷ) и модуля случайных отклонений |е|.

*Ранг* – это место данного числового значения среди упорядоченных значений анализируемого показателя.

Коэффициент ранговой корреляции рассчитывается по формуле (15):

 (15)

где - разность между рангами значений переменной *Х* и модуля *е*.

Необходимо проверить значимость коэффициента ранговой корреляции с помощью механизма проверки статистических гипотез.

Для этого формулируется гипотеза:

Н0: rx,e = 0 – коэффициент ранговой корреляции незначим

Н1: rx,e ≠ 0 – коэффициент ранговой корреляции значим

Проверка гипотезы осуществляется по критерию Стьюдента. Наблюдаемое значение критерия Стьюдента находится по формуле (16):

 (16)

Критерий T имеет число степеней свободы ν = n – 2.

По таблице критических точек Стьюдента [1,3,6] определяется значение *критической точки tкр*, которое зависит от следующих показателей:

- уровня значимости α, значение которого задается при проверке гипотезы;

- числа степеней свободы ν.

Чтобы сделать вывод о значимости коэффициента ранговой корреляции необходимо сравнить между собой рассчитанное значение критерия Т и значение критической точки tкр.

Если значение критерия Т попадает в *область принятия гипотезы* (-tкр < Т < tкр), можно сделать следующие выводы:

1. коэффициент ранговой корреляции признается незначимым;
2. в генеральной совокупности наблюдается гомоскедастичность, т.е. дисперсия случайных отклонений постоянна.

Если значение критерия Т попадает в *одну из критических областей* Т ≤ -tкр или Т ≥ tкр), выводы оказываются следующими:

1. коэффициент ранговой корреляции признается значимым;
2. в генеральной совокупности наблюдается гетероскедастичность, т.е. дисперсия случайных отклонений не является постоянной.

***Пример 10.***

Проверим с помощью теста ранговой корреляции Спирмена при уровне значимости 0,05 наличие гетероскедастичности в линейной регрессионной модели, построенной в примере 2.

Рассчитаем коэффициент ранговой корреляции по формуле (15). Промежуточные расчеты представлены в таблице 8.

Таблица 8 – Промежуточные расчеты для теста ранговой корреляции Спирмена

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *i* | *xi* | *R(xi)* | *ei* | *│ei│* | *R(│ei│)* | *di* | *di2* |
| 1 | 1,0 | 3 | -2,393 | 2,393 | 2 | 1 | 1 |
| 2 | 0,8 | 2 | -4,091 | 4,091 | 6 | -4 | 16 |
| 3 | 0,5 | 1 | -5,639 | 5,639 | 8 | -7 | 49 |
| 4 | 2,5 | 5 | 3,344 | 3,344 | 5 | 0 | 0 |
| 5 | 1,8 | 4 | 2,400 | 2,400 | 3 | 1 | 1 |
| 6 | 4,0 | 6 | 6,082 | 6,082 | 9 | -3 | 9 |
| 7 | 5,7 | 7 | 8,518 | 8,518 | 11 | -4 | 16 |
| 8 | 9,0 | 11 | -6,460 | 6,460 | 10 | 1 | 1 |
| 9 | 8,3 | 9 | -0,404 | 0,404 | 1 | 8 | 64 |
| 10 | 8,8 | 10 | -4,158 | 4,158 | 7 | 3 | 9 |
| 11 | 7,5 | 8 | 2,802 | 2,802 | 4 | 4 | 16 |
| *Сумма* |  |  |  |  |  |  | 182 |



Проверим значимость коэффициента ранговой корреляции при уровне значимости 0,1. Сформулируем гипотезу:

Н0: rx,e = 0 – коэффициент ранговой корреляции незначим,

Н1: rx,e ≠ 0 – коэффициент ранговой корреляции значим.

Рассчитаем наблюдаемое значение критерия Стьюдента по формуле (16):



Критерий T имеет число степеней свободы ν = n – 2 = 11 – 2 = 9.

Определим критическую точку tкр по таблице критических точек Стьюдента (таблицы критических точек можно найти в учебной литературе [1,3,6] и справочных изданиях по эконометрике, а также в раздаточном материале к курсовой работе на сайте moodle.nirhtu.ru):

при уровне значимости α = 0,05 и числе степеней свободы ν = 9 критическая точка tкр = 2,262.

Поскольку -tкр < Т < tкр (-2,262 < 0,526 < 2,262), можно сделать вывод о том, что коэффициент ранговой корреляции признается незначимым. Таким образом, в генеральной совокупности наблюдается гомоскедастичность, т.е. дисперсия случайных отклонений постоянна.

***3 Тест Голдфельда-Квандта для обнаружения гетероскедастичности***

Тест Голдфельда-Квандта проводится по следующему алгоритму:

1. вся совокупность наблюдений размерностью *п* упорядочивается по возрастанию значений фактора *Х* (или Ŷ для множественной регрессии);
2. упорядоченная совокупность делиться на 3 части размерностью *k, п – 2\*k, k* соответственно; при этом k определяется из пропорции: при *п =30, k=11*;
3. строятся отдельные уравнения регрессии для 1-ой и 3-ей частей выборки, и рассчитываются остаточные дисперсии для каждой из рассматриваемых частей по формулам (17) и (18):

, (17)

где  - квадрат значения случайного отклонения из первой части выборки;

m – число факторов в регрессионной модели *m* (в парной регрессионной модели *m = 1*).

, (18)

где  - квадрат значения случайного отклонения из третьей части выборки;

1. проверяется гипотеза о равенстве дисперсий двух совокупностей:

Н0: Sе32 = Sе12 – дисперсии двух частей выборки практически не отличаются друг от друга

Н1: Sе32 > Sе12 – дисперсия третьей части выборки значимо больше дисперсии первой части

Для проверки гипотезы используется критерий Фишера, наблюдаемое значение которого рассчитывается по формуле (19):

 (19)

Критерий F имеет распределение Фишера с числами степеней свободы *ν1=ν2=k-m-1*.

По таблице критических точек Фишера [1,3,6] определяется значение *критической точки fкр*, которое зависит от следующих показателей:

- уровня значимости α, значение которого задается при проверке гипотезы;

- числа степеней свободы ν1;

- числа степеней свободы ν2.

Необходимо сравнить между собой рассчитанное значение критерия F и значение критической точки fкр.

Если значение критерия F попадает в *область принятия гипотезы* (F < fкр), можно сделать следующие выводы:

1. разница между остаточными дисперсиями признается незначимой;
2. наблюдается гомоскедастичность.

Если значение критерия F попадает в *критическую область* (F ≥ fкр), выводы оказываются следующими:

1. разница между остаточными дисперсиями признается значимой;
2. наблюдается гетероскедастичность.

Если при расчете остаточных дисперсий оказывается, что Sе12 > Sе32 , то гипотеза преобразуется следующим образом:

Н0: Sе12 = Sе32

Н1: Sе12 > Sе32

Для проверки гипотезы используется критерий:

 (20)

т.е. значение критерия F всегда должно быть больше единицы.

***Пример 11.***

Проверим с помощью теста Голдфельда-Квандта при уровне значимости 0,05 наличие гетероскедастичности в линейной регрессионной модели, построенной в примере 2.

1) Упорядочим совокупность наблюдений по возрастанию значений фактора Х (см. таблицу 9):

Таблица 9 – Упорядоченные наблюдения

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *X* | 0,5 | 0,8 | 1 | 1,8 | 2,5 | 4 | 5,7 | 7,5 | 8,3 | 8,8 | 9 |
| *Y* | 3 | 5 | 7 | 13 | 15 | 20 | 25 | 22 | 20 | 17 | 15 |

2) Определим размерность частей, на которые необходимо разбить исследуемую совокупность. При *п =30 k=11*. В нашем случае *п =11*, тогда 

Принимаем, что размерность 1-ой части *k=4*, размерность 2-ой части *п – 2\*k=11-2·4=3*, размерность 3-ей части *k=4.*

3) Построим отдельные уравнения регрессии для 1-ой и 3-ей частей выборки, и рассчитаем остаточные дисперсии.

Рассмотрим 1-ую часть исследуемой совокупности (первые 4 столбца наблюдений из таблицы 9). Для построения уравнения регрессии используем формулы (4) метода наименьших квадратов. Результаты промежуточных расчетов приведены в таблице 10.



Таким образом, уравнение линейной регрессии, построенное по 1-ой части выборки, имеет вид:

*Ŷ = -0,957 + 7,763⋅Х*

Для расчета остаточной дисперсии необходимо дополнить таблицу 10: рассчитать прогнозируемые значения результирующего показателя ŷi, случайные отклонения ei и их квадраты.

Таблица 10 – Промежуточные вычисления для построения уравнения регрессии по 1-ой части выборки и определения остаточной дисперсии

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ***i*** | ***xi*** | ***yi*** | ***xi2*** | ***xiyi*** | ***ŷi*** | ***ei*** | ***ei2*** |
| 1 | 0,5 | 3 | 0,25 | 1,5 | 2,925 | 0,075 | 0,0057 |
| 2 | 0,8 | 5 | 0,64 | 4,0 | 5,253 | -0,253 | 0,0642 |
| 3 | 1,0 | 7 | 1,00 | 7,0 | 6,806 | 0,194 | 0,0377 |
| 4 | 1,8 | 13 | 3,24 | 23,4 | 13,016 | -0,016 | 0,0003 |
| *Сумма* | 4,1 | 28 | 5,13 | 35,9 | - | - | 0,1078 |
| *Среднее* | 1,025 | 7 | 1,283 | 8,975 | - | - | - |

Остаточную дисперсию рассчитаем по формуле (17):



Рассмотрим 3-ью часть исследуемой совокупности (последние 4 столбца наблюдений из таблицы 9). Для построения уравнения регрессии используем формулы (4) метода наименьших квадратов. Результаты промежуточных расчетов приведены в таблице 11.

Таблица 11 – Промежуточные вычисления для построения уравнения регрессии по 3-ей части выборки и определения остаточной дисперсии

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ***i*** | ***xi*** | ***yi*** | ***xi2*** | ***xiyi*** | ***ŷi*** | ***ei*** | ***ei2*** |
| 1 | 7,5 | 22 | 56,25 | 165,0 | 22,530 | -0,530 | 0,281 |
| 2 | 8,3 | 20 | 68,89 | 166,0 | 18,948 | 1,052 | 1,107 |
| 3 | 8,8 | 17 | 77,44 | 149,6 | 16,709 | 0,291 | 0,085 |
| 4 | 9,0 | 15 | 81,00 | 135,0 | 15,813 | -0,813 | 0,662 |
| *Сумма* | 33,6 | 74 | 283,58 | 615,6 | - | - | 2,134 |
| *Среднее* | 8,4 | 18,5 | 70,9 | 153,9 | - | - | - |



Таким образом, уравнение линейной регрессии, построенное по 3-ей части выборки, имеет вид:

*Ŷ = 56,112 – 4,478⋅Х*

Дополним таблицу 11 и рассчитаем остаточную дисперсию по формуле (18):



4) Проверим гипотезу о равенстве дисперсий двух совокупностей:

Н0: Sе32 = Sе12 – дисперсии двух частей выборки практически не отличаются друг от друга

Н1: Sе32 > Sе12 – дисперсия третьей части выборки значимо больше дисперсии первой части

Найдем расчетное значение *F*-критерия по формуле (19):



Критерий F имеет числа степеней свободы ν1 = ν2 = k – m – 1 = 4 – 1 – 1 = 2.

Определим критическую точку fкр по таблице критических точек Фишера (таблицы критических точек можно найти в учебной литературе [1,3,6] и справочных изданиях по эконометрике, а также в раздаточном материале к курсовой работе на сайте moodle.nirhtu.ru):

при уровне значимости α = 0,05 и числах степеней свободы ν1 = 2 и ν2 = 2 критическая точка fкр = 19.

Поскольку F > fкр (19,796 > 19), можно сделать вывод о том, что разница между остаточными дисперсиями признается значимой, и в построенном уравнении регрессии наблюдается гетероскедастичность. Однако значение критерия F лишь незначительно превышает критическую точку fкр, поэтому гетероскедастичность в данном случае незначительная, и ею можно пренебречь.

При обнаружении значительной гетероскедастичности необходимо выяснить причины ее возникновения и предложить способ устранения.

Основными *причинами наличия гетероскедастичности* в регрессионных моделях являются:

* ошибки спецификации – влияние значимого фактора, не учтенного в модели; для устранения гетероскедастичности в этом случае необходимо изменить спецификацию модели (подробнее см. п. 4.5.1);
* внутренние свойства изучаемых переменных (например, при анализе зависимости спроса от дохода потребителя выясняется, что чем больше доход, тем больше индивидуальное значение спроса колеблется относительно ожидаемого значения); в этом случае необходимо использовать специальные методы построения эконометрической модели.

## 4.5 Изменение спецификации модели

## 4.5.1. Спецификация регрессионных моделей

Целью эконометрики является построение такой математической модели, которая наилучшим образом описывает изучаемое экономическое явление.

Для этого уже на начальном этапе анализа необходимо сформулировать как можно больше исходных представлений о взаимосвязях между рассматриваемыми переменными на основе экономической теории. Полученная информация позволяет выбрать необходимые факторы и математическую функцию для построения математической модели.

Например, при построении модели производственной функции (зависимости объема выпуска продукции от объема производственных ресурсов) в качестве влияющих факторов могут выступать:

- стоимость материальных запасов;

- фонд оплаты труда;

- стоимость основных производственных фондов;

- объем капитальных вложений в улучшение производственных мощностей и др.

Выбор конкретных факторов зависит от цели исследования, а также от возможностей проведения статистического анализа (имеется ли вся необходимая информация, есть ли возможность собрать дополнительные данные, каковы затраты на сбор и обработку дополнительных данных и т.п.).

Результатом анализа на данном этапе является ***спецификация эконометрической модели*** – определение модели в виде уравнения с указанием рассматриваемых переменных (факторов и результирующего показателя), функциональной формы связи между переменными, а также априорных ограничений на параметры модели (например, производственная функция должна быть возрастающей, т.е. угловой коэффициент должен быть положительным).

Предположим, получена начальная спецификация производственной функции в виде линейной модели с одним фактором – стоимостью материальных запасов:

M(Y|X1,X2) = β0 + β1 \* Х

После этого на основе имеющихся эмпирических данных (выборки) производится построение эмпирической модели, т.е. оцениваются эмпирические коэффициенты b0 и b1. Для этого, как правило, используется метод наименьших квадратов. В результате строится эмпирическое уравнение регрессии:

Ŷ = b0 + b1 \* Х

Затем производится проверка адекватности модели имеющимся эмпирическим данным, а также проверяются различные гипотезы о значениях параметров модели (о значимости коэффициентов и др.). Если модель имеет невысокое качество, говорят, что ***модель неправильно специфицирована***, или что имеются ***ошибки спецификации***. В этом случае необходимо устранить выявленные недостатки, т.е. предложить *новую спецификацию модели*, например:

- добавить новый фактор;

- устранить фактор, не оказывающий сильного влияния на модель;

- изменить функциональную форму модели;

- рассмотреть только часть имеющихся статистических данных (если обнаружены выбросы или кластеринг) и др.

После выбора новой спецификации повторяются этапы построения эмпирического уравнения и проверки его качества. Так происходит до тех пор, пока не будет построена удовлетворительная эконометрическая модель, которую можно использовать для анализа и прогнозирования.

## 4.5.2 Ошибки спецификации

На начальных этапах построения уравнений регрессии довольно часто возникают ошибки спецификации. Они появляются обычно из-за поверхностных знаний об исследуемых экономических процессах, или из-за погрешностей сбора и обработки статистических данных. Важно уметь обнаружить и исправить эти ошибки.

Можно привести следующие рекомендации по обнаружению ошибок спецификации и совершенствованию регрессионных моделей:

1. Если один или несколько (но не все) угловые коэффициенты в уравнении регрессии оказываются незначимыми, это говорит о том, что соответствующие факторы не влияют на результирующий показатель, и их можно исключить из модели без потери качества.

Однако если незначим свободный член, его не рекомендуется исключать из модели, так как он не связан с факторами, и только уточняет расположение линии регрессии на поле корреляции. Незначимость свободного члена не является ошибкой спецификации.

1. Если уравнение имеет низкий коэффициент детерминации, это может иметь две причины: в модели отсутствует какой-то важный фактор, и его необходимо добавить; либо неверно подобрана математическая функция, лежащая в основе модели, и ее нужно изменить.
2. Если обнаружена автокорреляция остатков, это может говорить о том, что неверно подобрана математическая функция, лежащая в основе модели.
3. Если обнаружена гетероскедастичность, это может быть вызвано либо проблемами в наборе данных (выбросы, кластеринг), либо неправильным выбором факторов.

При обнаружении ошибок спецификации, необходимо внести предложения по изменению спецификации модели.

***Пример 12***

Сведем вместе результаты проверки качества линейной регрессионной модели, построенной в примере 2 и сделаем выводы о наличии ошибок спецификации.

При проверке качества получены следующие результаты (см. примеры 3 – 11):

- значение коэффициента детерминации R2 = 0,518 показывает, что 51,8% разброса значений прибыли предприятия объясняется с помощью уравнения регрессии (изменением расходов на рекламу), а 48,2% объясняется случайными причинами; таким образом, построенное уравнение может быть признано адекватным эмпирическим данным, однако адекватность уравнения достаточно низкая, так как коэффициент детерминации намного меньше 1;

- коэффициент детерминации R2 в генеральной совокупности является значимым, т.е. в генеральной совокупности существует значимая линейная связь между прибылью предприятия и расходами на рекламу, и построенное уравнение регрессии адекватно данным генеральной совокупности;

- теоретический угловой коэффициент регрессии и теоретический свободный член регрессии являются значимыми, т.е. в генеральной совокупности существует значимая линейная связь между прибылью предприятия и расходами на рекламу, и линия регрессии не может проходить через начало координат;

- наблюдается положительная автокорреляция остатков в регрессионной модели, т.е. не выполняется предпосылка МНК о некоррелированности случайных отклонений εi;

- тест ранговой корреляции Спирмена показал, что в генеральной совокупности наблюдается гомоскедастичность, т.е. дисперсия случайных отклонений постоянна; тест Голдфельда-Квандта показал наличие незначительной гетероскедастичности, которой можно пренебречь; считаем, что предпосылка МНК о постоянстве дисперсии случайных отклонений εi выполняется.

Рассмотрим возможные ошибки спецификации:

Коэффициент детерминации R2=0,518 показывает, что адекватность уравнения достаточно низкая, так как коэффициент детерминации намного меньше 1. Так как при этом обнаружена положительная автокорреляция остатков, то можно предположить неправильно выбранную форму модели. В этом необходимо убедиться по виду поля корреляции.

В примере 1 на поле корреляции выявлена нелинейная взаимосвязь, которую можно моделировать в виде параболы (полиномиальной функции второго порядка).

Вывод: чтобы повысить качество модели и устранить ошибки спецификации необходимо изменить форму модели на полиномиальной функции второго порядка.

Новая спецификация модели: Ŷ = b0 + b1 \* Х + b2 \* Х2

***Пример 13***

По выборке объемом 40 оценена следующая модель производственной функции:

Ŷ = 0,6 + 0,46 \* L + 0,32 \* K,

где Y – объем выпуска;

 L – затраты труда;

 К – затраты капитала (основных и оборотных фондов).

Предположим, что при проверке качества получены следующие результаты:

- значение коэффициента детерминации R2 = 0,65 показывает, что 65% разброса значений объема выпуска объясняется с помощью уравнения регрессии (изменением затрат труда), а 35% объясняется случайными причинами; таким образом, построенное уравнение может быть признано адекватным эмпирическим данным;

- теоретический свободный член регрессии β0 значим, следовательно, линия регрессии не может проходить через начало координат;

- теоретический угловой коэффициент β1 незначим, т.е. фактор L (затраты труда) не влияет на объем выпуска;

- теоретический угловой коэффициент β2 значим, т.е. в генеральной совокупности существует значимая линейная связь между затратами капитала и объемом выпуска;

- автокорреляция остатков в регрессионной модели отсутствует, т.е. выполняется предпосылка МНК о некоррелированности случайных отклонений εi.

Рассмотрим возможные ошибки спецификации:

1. Один из угловых коэффициентов незначим, т.е. соответствующий фактор (затраты труда) не влияет на результирующий показатель. Указанный фактор можно исключить из уравнения регрессии без потери качества.

2. Значение коэффициента детерминации R2 объясняет менее 90% разброса результирующего показателя, т.е. есть возможность повысить адекватность уравнения. В построенной модели отсутствует автокорреляция остатков, следовательно, математическая функция подобрана верно. Значит, можно рекомендовать увеличить адекватность модели, добавив в нее какой-нибудь значимый фактор.

*Общий вывод*: чтобы повысить качество модели и устранить ошибки спецификации необходимо исключить фактор L и заменить его на другой влияющий фактор.

В производственных функциях, кроме факторов трудозатрат и затрат капитала, часто рассматривается фактор I – инвестиции в производство. Можно рекомендовать рассмотреть этот фактор.

Новая спецификация модели:

Ŷ = b0 + b1 \* K + b2 \* I

## 4.5.3 Построение множественной регрессионной модели

Прогнозирование одной переменной Y на основании нескольких факторов Х1, Х2, …, Хm производится с помощью уравнения ***множественной регрессии***.

Эмпирическое уравнение множественной линейной регрессии для *m* факторов выглядит следующим образом:

Ŷ = b0 + b1 · X1 + b2 · X2 + … + bm · Xm

Угловые коэффициенты bj показывают изменение результирующего показателя Y при изменении соответствующего фактора Хj на единицу при условии, что все остальные факторы остаются неизменными.

Свободный член b0 показывает значение результирующего показателя Y при нулевых значениях всех факторов.

Чаще всего для определения коэффициентов уравнения регрессии используется метод наименьших квадратов (МНК), который минимизирует сумму квадратов отклонений точек наблюдений от линии регрессии.

Вычисление коэффициентов множественной регрессии вручную является достаточно трудоемким процессом, причем трудоемкость многократно возрастает с увеличением числа факторов. Поэтому построение уравнения множественной регрессии обычно производят с помощью различных компьютерных программ.

При выполнении курсовой работы рекомендуется определять коэффициенты уравнения регрессии с помощью функции ЛИНЕЙН (см. Приложение В).

Проверка адекватности эмпирического уравнения множественной регрессии данным генеральной совокупности осуществляется с помощью проверки значимости коэффициента детерминации. Цель этой проверки заключается в том, чтобы выяснить, *объясняет ли совместное влияние всех m факторов значимую долю разброса результирующего показателя Y*.

Нулевая и альтернативная гипотезы формулируются следующим образом:

Н0: R2 = 0 – значение коэффициента детерминации незначимо, совместное влияние *m* факторов объясняет незначительную долю разброса переменной Y.

Н1: R2 > 0 – значение коэффициента детерминации значимо, совместное влияние *m* факторов объясняет значительную долю разброса переменной Y.

Для проверки гипотезы используется критерий Фишера:



Критерий F имеет распределение Фишера с числами степеней свободы ν1 = m и ν2 = n-m-1.

Если *коэффициент детерминации незначим*, т.е. совместное влияние *m* факторов объясняет незначимую долю разброса переменной Y, *уравнение регрессии сразу должно быть признано некачественным*. Дальнейшая проверка его качества не имеет смысла.

Если же *коэффициент детерминации оказывается значимым*, т.е. *совместное* влияние *m* факторов объясняет значимую долю разброса переменной Y, уравнение регрессии признается адекватным данным генеральной совокупности.

При этом неизвестно, какие из факторов действительно влияют на результат, а какие нет. В этом случае можно продолжить проверку качества уравнения регрессии, используя гипотезы о значимости коэффициентов регрессии.

Проверка значимости каждого из угловых коэффициентов регрессии βj показывает, насколько значимым является влияние соответствующего фактора Хj на Y при условии, что все остальные факторы остаются неизменными.

Если выяснить, какие из используемых факторов не влияют на результат, их можно исключить из уравнения регрессии без снижения его качества. При этом уравнение может значительно упроститься.

## 4.5.4 Мультиколлинеарность

Чтобы по методу наименьших квадратов (МНК) можно было получить наилучшие оценки коэффициентов регрессии, необходимо, чтобы выполнялись предпосылки МНК (условия Гаусса-Маркова) [1, 2, 4, 6].

Для множественной регрессии одной из предпосылок является отсутствие ***мультиколлинеарности*** – сильной линейной зависимости между какими-либо факторами, участвующими в уравнении.

Признаком наличия мультиколлинеарности является следующее несоответствие выводов: все угловые коэффициенты в уравнении регрессии незначимы, и при этом коэффициент детерминации имеет высокое значение.

Для устранения мультиколлинеарности необходимо исключить один из взаимосвязанных факторов.

## 4.6 Сравнение качества моделей

В некоторых случаях могут быть построены несколько эконометрических моделей удовлетворительного качества (в курсовой работе строятся две регрессионные модели). Для выбора наилучшей модели следует сравнить построенные модели по следующим критериям:

1. *Адекватность* (сравниваются значения коэффициента детерминации).
2. *Согласованность с теорией* (определяется соответствие параметров модели сформулированным начальным ограничениям).
3. *Простота* (выбирается модель с наименьшим количеством факторов; при сравнении линейной и нелинейной модели выбирается линейная модель).

При сравнении моделей, имеющих разное количество факторов (например, парной и множественной модели), по критерию адекватности может возникнуть ситуация, когда коэффициенты детерминации отличаются незначительно.

Коэффициент детерминации R2 обладает следующим свойством: при увеличении количества факторов в модели он всегда возрастает, даже если новый фактор не является значимым для модели.

Поэтому обоснованный вывод можно сделать на основе проверки следующей гипотезы:

Н0: R12 = R22 – качество двух уравнений практически не отличается;

Н1: R12 > R22 – первое уравнение (с большим количеством факторов) качественнее.

Для проверки гипотезы используется критерий:

,

где R12 – коэффициент детерминации первого уравнения (с большим количеством факторов);

 R22 – коэффициент детерминации второго уравнения (с меньшим количеством факторов);

 m – количество факторов в первом уравнении;

 k – количество факторов, исключаемых из второго уравнения (по сравнению с первым).

Критерий F подчинен распределению Фишера с числами степеней свободы ν1 = k и
ν2 = n-m-1.

Если принимается нулевая гипотеза, качество двух уравнений по критерию адекватности практически не отличается, поэтому для моделирования можно использовать более простое уравнение – с меньшим количеством факторов.

Если принимается альтернативная гипотеза, уравнение с большим количеством факторов значимо более адекватно, чем более простое уравнение. Поэтому для моделирования лучше использовать уравнение с большим количеством факторов.

***Пример 14***

По выборке объемом 15 оценены следующие модели производственной функции:

Ŷ = 0,85 + 0,33 \* L + 0,25 \* K (R12 = 0,76)

Ŷ = 3,44 + 0,51 \* K (R22 = 0,68)

где Y – объем выпуска;

 L – затраты труда;

 К – затраты капитала (основных и оборотных фондов).

Сравним качество двух моделей по нескольким критериям качества:

1. По критерию *адекватности* более предпочтительным является уравнение множественной регрессии, так как R12 > R22.

2. По критерию *согласованности с теорией* оба уравнения соответствуют экономической теории, так как угловые коэффициенты больше нуля, т.е. с увеличением объема ресурса объем выпуска продукции увеличивается.

3. По критерию *простоты* более предпочтительным является уравнение парной регрессии, так как оно содержит меньше факторов.

Если показатели адекватности уравнений отличаются незначительно, рекомендуется выбрать более простое уравнение.

Если же показатели адекватности значительно отличаются, рекомендуется выбрать более адекватное уравнение.

Адекватность уравнений с разным количеством факторов сравним с помощью следующей гипотезы:

Н0: R12 = R22 – качество двух уравнений практически не отличается;

Н1: R12 > R22 – первое уравнение (с большим количеством факторов) качественнее.

Для проверки гипотезы используется критерий:

,

Рассчитаем критерий Фишера:



Определим критическую точку распределения Фишера при уровне значимости 0,05 и числах степеней свободы ν1 = 1 и ν2 = 12: fкр = 4,74.

Значение критерия F попало в область принятия гипотезы, следовательно, принимается нулевая гипотеза, т.е. качество двух уравнений по критерию адекватности практически не отличается. Значит, для моделирования можно использовать более простое уравнение – с меньшим количеством факторов: Ŷ = 3,44 + 0,51 \* K.

# 5 Задания на курсовую работу

## Вариант 1

*Теоретический вопрос:*

Определение эконометрики. Взаимосвязь эконометрики с другими науками. Цели эконометрики.

*Практическое задание:*

Исследуется зависимость доли расходов на товары длительного пользования (*Y*, %) от дохода семьи (*Х*, тыс. руб.):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 10,048 | 16,231 | 11,865 | 17,611 | 14,410 | 11,341 | 9,203 | 18,407 | 7,864 | 17,038 |
| **X** | 11 | 32 | 13 | 45 | 20 | 18 | 12 | 55 | 10 | 60 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 16,524 | 11,309 | 17,529 | 14,129 | 18,771 | 11,360 | 16,935 | 16,931 | 17,610 | 12,615 |
| **X** | 24 | 9 | 58 | 25 | 50 | 14 | 36 | 28 | 40 | 16 |

Задание:

1. Исследуйте корреляционную взаимосвязь между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ* *= b0 + b1 X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,1;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,1;
* проверьте наличие автокорреляции остатков графическим методом и с помощью теста Дарбина-Уотсона при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,01;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Подготовьте исходные данные для построения логарифмической регрессионной модели и рассчитайте ее параметры.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,1;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,1;
* проверьте наличие автокорреляции остатков графическим методом и с помощью теста Дарбина-Уотсона при уровне значимости 0,05.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ые доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения доли расходов на товары длительного пользования при доходах семьи 57 тыс. руб. и 30 тыс. руб.

## Вариант 2

*Теоретический вопрос:*

Характер зависимости между двумя переменными. Корреляционная зависимость между двумя переменными.

*Практическое задание:*

По имеющимся статистическим данным исследуется кривая Филипса – зависимость между темпом прироста заработной платы W (%) и уровнем безработицы U:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **U** | 1,38 | 1,15 | 1,50 | 1,55 | 1,20 | 1,10 | 1,00 | 1,35 | 1,05 | 1,80 |
| **W** | 3,012 | 7,778 | 4,306 | 4,878 | 3,302 | 5,778 | 8,163 | 3,922 | 7,018 | 4,494 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **U** | 1,90 | 1,45 | 1,85 | 1,20 | 1,50 | 1,25 | 1,40 | 1,30 | 1,60 | 1,65 |
| **W** | 2,198 | 3,5 | 4,437 | 4,98 | 4,127 | 3,67 | 4,127 | 4,167 | 5,263 | 4,003 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *W* и *U*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŵ = b0 + b1U* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,1;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,1;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,05;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Подготовьте исходные данные для построения гиперболической регрессионной модели и рассчитайте ее параметры.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,1;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,1;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 99%-ые доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы средних значений темпа прироста заработной платы, если уровень безработицы составит 2 и 1,48.

## Вариант 3

*Теоретический вопрос:*

Числовые характеристики, используемые для измерения силы взаимосвязи между переменными.

*Практическое задание:*

Анализируется зависимость выпуска продукции (Y, тыс. ед.) от количества израсходованных материалов (Х, кг) в некоторой отрасли. Получена следующая выборка:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Выпуск** | 100 | 220 | 180 | 470 | 520 | 590 | 700 | 150 | 120 | 750 |
| **Материалы** | 900 | 6400 | 6000 | 25600 | 52000 | 74000 | 120000 | 2000 | 1200 | 143400 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Выпуск** | 410 | 470 | 310 | 250 | 130 | 160 | 240 | 110 | 200 | 730 |
| **Материалы** | 29000 | 50000 | 15200 | 8400 | 1500 | 2700 | 7700 | 1100 | 4800 | 132600 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфелда-Квандта при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Подготовьте исходные данные для построения степенной регрессионной модели и рассчитайте ее параметры.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 95%-ые доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения выпуска продукции при потреблении 1000 кг и 70 000 кг материала.

## Вариант 4

*Теоретический вопрос:*

Детерминированные и стохастические экономические модели.

*Практическое задание:*

Анализируется прибыль предприятия (Y, млн. руб.) в зависимости от расходов на рекламу (Х, млн. руб.). Получена следующая выборка:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **X** | 0,8 | 1 | 1,8 | 2,5 | 4 | 5,7 | 7,5 | 8,3 | 8,8 | 0,7 |
| **Y** | 5 | 7 | 16 | 15 | 20 | 27 | 25 | 20 | 17 | 8 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **X** | 1,4 | 3,1 | 2,2 | 3 | 2,8 | 3,9 | 5 | 4,3 | 6,3 |
| **Y** | 11 | 16 | 18 | 17 | 16 | 19 | 24 | 21 | 24 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,05;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Подготовьте исходные данные для построения полиномиальной регрессионной модели второго порядка и рассчитайте ее параметры.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,01.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 99%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы средних значений прибыли предприятия, если расходы на рекламу составляют 6 млн. руб. и 2 млн. руб.

## Вариант 5

*Теоретический вопрос:*

Функция регрессии и регрессионная модель.

*Практическое задание:*

Изучается зависимость выпуска продукции (*Y*, тыс. ед.) от ее материалоемкости (*Х*, кг на единицу продукции) в некоторой отрасли. Имеется следующая выборка:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Материалоемкость** | 9 | 6 | 5 | 4 | 3,7 | 3,6 | 3,5 | 6 | 7 | 3,4 |
| **Выпуск** | 100 | 220 | 230 | 400 | 520 | 590 | 700 | 150 | 120 | 750 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Материалоемкость** | 4,2 | 4,8 | 5,1 | 5,4 | 7,0 | 6,6 | 5,6 | 8,1 | 6,0 |
| **Выпуск** | 410 | 470 | 310 | 250 | 130 | 160 | 240 | 110 | 200 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфелда-Квандта при уровне значимости 0,05;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Подготовьте исходные данные для построения гиперболической регрессионной модели и рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,01.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения выпуска при потреблении 3 кг и 6 кг материала на единицу продукции.

## Вариант 6

*Теоретический вопрос:*

Причины наличия в регрессионной модели случайного отклонения.

*Практическое задание:*

Исследуется зависимость между возрастом рабочих (Х, лет) и их среднедневной выработкой (Y, руб.) по следующей выборке:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Возраст** | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 22 |
| **Выработка** | 50 | 60 | 80 | 80 | 100 | 85 | 60 | 50 | 55 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Возраст** | 50 | 24 | 48 | 32 | 40 | 35 | 38 | 41 | 39 |
| **Выработка** | 70 | 55 | 60 | 87 | 90 | 96 | 92 | 98 | 97 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,1;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,1;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Подготовьте исходные данные для построения полиномиальной регрессионной модели второго порядка и рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,1;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,1;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения выработки рабочих 19 и 45 лет.

## Вариант 7

*Теоретический вопрос:*

Теоретическое и эмпирическое уравнение регрессии.

*Практическое задание:*

По 20 территориям Северного, Северо-Западного и Центрального районов известны данные о зависимости потребительских расходов на душу населения (*Y*, тыс. руб.) от денежных доходов на душу населения (*Х*, тыс. руб.) за ноябрь 1997 г.:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 596 | 417 | 354 | 526 | 934 | 412 | 525 | 367 | 364 | 336 |
| **X** | 913 | 1095 | 606 | 876 | 1314 | 593 | 754 | 528 | 520 | 539 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 409 | 452 | 367 | 328 | 460 | 380 | 439 | 344 | 401 | 514 |
| **X** | 540 | 682 | 537 | 589 | 626 | 521 | 626 | 521 | 658 | 746 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,1;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,1;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,05;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Устраните проблемы в наборе данных, найденные в п. 1. Предложите функциональную форму регрессионной модели на основании анализа поля корреляции. Рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,1;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,1;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,05.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения потребительских расходов на душу населения при денежном доходе 1200 тыс. руб. и 800 тыс. руб.

## Вариант 8

*Теоретический вопрос:*

Использование метода наименьших квадратов для расчета коэффициентов парной регрессии.

*Практическое задание:*

По 14 территориям Уральского и Западно-Сибирского районов известны данные о зависимости потребительских расходов на душу населения (*Y*, тыс. руб.) от денежных доходов на душу населения (*Х*, тыс. руб.) за ноябрь 1997 г.:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 461 | 424 | 298 | 351 | 524 | 484 | 425 |
| **X** | 912 | 809 | 748 | 847 | 1087 | 1074 | 1008 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 277 | 321 | 573 | 576 | 388 | 497 | 563 |
| **X** | 682 | 697 | 1951 | 967 | 898 | 1963 | 2027 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Устраните проблемы в наборе данных, найденные в п. 1. Предложите функциональную форму регрессионной модели на основании анализа поля корреляции. Рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 99%-ые доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения потребительских расходов на душу населения при денежном доходе 1500 тыс. руб. и 1100 тыс. руб.

## Вариант 9

*Теоретический вопрос:*

Предпосылки метода наименьших квадратов (условия Гаусса-Маркова) для парной регрессии.

*Практическое задание:*

По 20 предприятиям некоторой отрасли анализируют среднедневную заработную плату сотрудников (*Y*, руб.) в зависимости от масштаба (количества сотрудников) предприятия (*Х*, чел.):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 100 | 90 | 85 | 110 | 130 | 165 | 190 | 200 | 220 | 215 |
| **Y** | 750 | 828 | 801 | 812 | 975 | 968 | 1003 | 1185 | 1115 | 997 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 275 | 290 | 315 | 330 | 400 | 395 | 425 | 460 | 500 | 510 |
| **Y** | 1345 | 1250 | 1516 | 1247 | 1458 | 1639 | 1413 | 1458 | 1621 | 2694 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Устраните проблемы в наборе данных, найденные в п. 1. Предложите функциональную форму регрессионной модели на основании анализа поля корреляции. Рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднедневной заработной платы сотрудников предприятий численностью 415 и 600 чел.

## Вариант 10

*Теоретический вопрос:*

Спецификация регрессионной модели. Ошибки спецификации.

*Практическое задание:*

Выдвигается предположение, что средняя заработная плата (*Y,* руб.) рабочих пропорциональна их стажу (*X*, лет). Для анализа данного утверждения обследуются 20 рабочих:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **x** | 3 | 1 | 2 | 6 | 10 | 9 | 12 | 16 | 19 | 18 |
| **y** | 17000 | 13000 | 13000 | 20000 | 21000 | 19000 | 21000 | 25000 | 23000 | 22000 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **x** | 21 | 25 | 24 | 25 | 27 | 32 | 29 | 36 | 38 | 40 |
| **y** | 26000 | 32000 | 21000 | 28000 | 26000 | 33000 | 34000 | 31500 | 16500 | 37500 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,01;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Устраните проблемы в наборе данных, найденные в п. 1. Предложите функциональную форму регрессионной модели на основании анализа поля корреляции. Рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,01.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы средней заработной платы рабочих, имеющих стаж работы 15 и 35 лет.

## Вариант 11

*Теоретический вопрос:*

Гетероскедастичность (суть, последствия, методы обнаружения).

*Практическое задание:*

Исследуется зависимость между доходом (*Х*, тыс. руб.) домохозяйств и их расходом (*Y*, тыс. руб.) на продукты питания:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **X** | 20,4 | 21,76 | 28,56 | 31,68 | 32 | 33,28 | 35,36 | 36,4 | 38,64 | 41,84 |
| **Y** | 11,6 | 9,04 | 10,8 | 9,92 | 8,24 | 11,12 | 9,28 | 13,2 | 11,04 | 14,56 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **X** | 44,56 | 48,96 | 50 | 55,76 | 59,04 | 60,64 | 63,36 | 65,92 | 66,4 | 69,12 |
| **Y** | 13,04 | 8,72 | 12,88 | 23,2 | 11,44 | 24,88 | 15,84 | 23,6 | 13,84 | 10,96 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,1;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,1;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Устраните проблемы в наборе данных, найденные в п. 1. Предложите функциональную форму регрессионной модели на основании анализа поля корреляции. Рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,1;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,1;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 95%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения расходов на продукты питания домохозяйств с доходом 25 и 48 тыс. руб.

## Вариант 12

*Теоретический вопрос:*

Гетероскедастичность (суть, последствия, методы устранения).

*Практическое задание:*

Проводится анализ зависимости средней заработной платы (Y, руб.) от средней производительности (Х, руб.) на предприятиях различного масштаба (численности сотрудников):

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **X** | 9320 | 8630 | 8050 | 8320 | 8600 | 9120 | 9540 | 9730 | 10120 | 10740 | 11200 |
| **Y** | 3320 | 3640 | 3900 | 4120 | 4090 | 4200 | 4380 | 4500 | 4610 | 4800 | 5000 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Устраните проблемы в наборе данных, найденные в п. 1. Предложите функциональную форму регрессионной модели на основании анализа поля корреляции. Рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,1.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ые доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы средней заработной платы на предприятиях со средней производительностью 15 и 21тыс. руб.

## Вариант 13

*Теоретический вопрос:*

Сезонная корректировка регрессионной модели.

*Практическое задание:*

В целях прогнозирования объема экспорта страны на будущие периоды были собраны данные за 25 лет по объему экспорта (*Y*, млрд. долл.) и индексу физического объема промышленного производства (*Х*, % к предыдущему году):

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 37,13 | 30,21 | 17,64 | 9,88 | 54,91 | 29,93 | 51,86 | 34,69 |
| **Y** | 14,49 | 3,24 | 16,66 | 2,39 | 21,84 | 22,18 | 19,02 | 32,69 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 71,41 | 65,77 | 65,24 | 83,01 | 74,83 | 86,69 | 94,03 | 92,61 |
| **Y** | 28,02 | 38,58 | 28,50 | 46,63 | 11,79 | 45,10 | 45,74 | 43,36 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 106,99 | 86,50 | 90,72 | 109,47 | 96,28 | 100,92 | 147,57 | 150,53 | 121,97 |
| **Y** | 39,31 | 26,21 | 52,02 | 26,54 | 53,48 | 33,52 | 70,42 | 45,90 | 76,57 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Подготовьте исходные данные для построения степенной регрессионной модели и рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 95%-ые доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения объема экспорта при индексах физического объема промышленного производства 130% и 85%.

## Вариант 14

*Теоретический вопрос:*

Автокорреляция остатков (суть, причины, последствия, методы обнаружения).

*Практическое задание:*

Изучается зависимость объема продаж бензина (Y, тыс. л) от индекса потребительских цен (Х, % к базовому кварталу). Полученные за последние 20 кварталов данные представлены в таблице:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 98,57 | 94,69 | 95,42 | 83,2 | 120,11 | 116,19 | 114,99 | 115,98 | 111,43 | 136,49 |
| **Y** | 93,83 | 97,18 | 78,63 | 85,69 | 90,43 | 87,98 | 76,3 | 79,97 | 70,16 | 64,31 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 140,26 | 127,76 | 123,38 | 129,74 | 147,49 | 146,46 | 141,44 | 131,77 | 158,28 | 138,06 |
| **Y** | 69,12 | 69,4 | 66,36 | 74,5 | 67,96 | 65,18 | 65,38 | 69,04 | 61,23 | 69,78 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Подготовьте исходные данные для построения экспоненциальной регрессионной модели и рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 95%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения объема продаж бензина при индексах потребительских цен 160% и 120%.

## Вариант 15

*Теоретический вопрос:*

Автокорреляция остатков (суть, причины, последствия, методы устранения).

*Практическое задание:*

Изучается годовое потребление товара А (*Y*, тыс. руб.) в зависимости от доходов населения (*Х*, тыс. руб.) по данным 20 лет:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 31,1 | 43,73 | 50,9 | 62,35 | 49,29 | 63,4 | 74,05 | 61,9 | 66,14 | 87,1 |
| **Y** | 51,68 | 50,49 | 55,61 | 56,26 | 65,9 | 57,46 | 50,22 | 70,56 | 59,76 | 71,92 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 83,52 | 104,14 | 84,16 | 108,63 | 94,13 | 117,33 | 101,89 | 123,61 | 114,45 | 130,95 |
| **Y** | 54,99 | 66,13 | 70,95 | 79,14 | 58,59 | 88,97 | 80,03 | 108,74 | 102,59 | 114,94 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Подготовьте исходные данные для построения полиномиальной регрессионной модели второго порядка и рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,1.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения объема потребления товара А при доходах населения 90 тыс. руб. и 140 тыс. руб.

## Вариант 16

*Теоретический вопрос:*

Моделирование структурных изменений в регрессионной модели. Тест Чоу.

*Практическое задание:*

Изучается зависимость между уровнем дивидендов, выплачиваемых фирмой по обыкновенным акциям (*Y*, %), и среднегодовой стоимостью основных фондов фирмы (*Х*, млн. руб.) за последние 20 лет:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 73,81 | 74,22 | 74,52 | 75,18 | 75,34 | 76,06 | 76,4 | 76,63 | 76,98 | 77,4 |
| **Y** | 4,79 | 3,97 | 3,59 | 3,62 | 2,3 | 1,9 | 2,69 | 2,93 | 2,43 | 1,98 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 77,97 | 78,14 | 78,79 | 79 | 79,81 | 79,86 | 80,67 | 80,64 | 81,2 | 81,65 |
| **Y** | 1,94 | 2,26 | 1,4 | 1,79 | 1,27 | 1,58 | 1,41 | 1,12 | 0,94 | 1,01 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Подготовьте исходные данные для построения степенной регрессионной модели и рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 99%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения уровня дивидендов при стоимости основных фондов 82 млн. руб. и 79 млн. руб.

## Вариант 17

*Теоретический вопрос:*

Основные элементы временного ряда. Моделирование тенденции временного ряда.

*Практическое задание:*

В таблице приведены данные за 20 лет о потреблении (*Y*, тыс. долл.) и личных доходах
(*Х*, тыс. долл.) населения:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 347,95 | 330,57 | 370,17 | 400,46 | 418,17 | 396,75 | 434,8 | 470,58 | 469,46 | 470,41 |
| **Y** | 367,84 | 414,35 | 462,41 | 452,85 | 463,22 | 410,43 | 422,09 | 464,55 | 400,21 | 465,38 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 492,26 | 531,62 | 512,48 | 555,61 | 553,06 | 592,5 | 607,6 | 616,56 | 631,43 | 658,32 |
| **Y** | 427,95 | 438,65 | 491,97 | 605,53 | 465,75 | 527,93 | 510,81 | 679,12 | 522,63 | 691,95 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,1;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,1;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Подготовьте исходные данные для построения экспоненциальной регрессионной модели и рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,1;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,1;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения потребления при доходах населения 680 тыс. руб. и 510 тыс. руб.

## Вариант 18

*Теоретический вопрос:*

Основные элементы временного ряда. Моделирование сезонных и циклических колебаний.

*Практическое задание:*

Имеются данные за 20 месяцев о производительности труда на одного рабочего (*Y*, усл. ед.) и электровооруженности труда (*Х*, кВт⋅ч/чел.-ч) на одном из промышленных предприятий:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 3,84 | 3,86 | 3,8 | 3,84 | 3,89 | 3,96 | 3,95 | 3,98 | 4,05 | 4,06 |
| **Y** | 36,24 | 34,02 | 37,42 | 33,9 | 34,63 | 36,52 | 36,05 | 39,21 | 34,6 | 39,13 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Х** | 4,01 | 4,05 | 4,15 | 4,15 | 4,16 | 4,2 | 4,16 | 4,21 | 4,29 | 4,23 |
| **Y** | 36,89 | 40,68 | 42,51 | 42,15 | 45,99 | 45,68 | 45,96 | 46,95 | 47,95 | 48,49 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Подготовьте исходные данные для построения полиномиальной регрессионной модели второго порядка и рассчитайте параметры нового уравнения.
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,05.
1. Нанесите на поле корреляции графики двух функций регрессии. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 95%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения производительности труда при значениях электровооруженности 4,35 кВт⋅ч/чел.-ч и 3,9 кВт⋅ч/чел.-ч.

## Вариант 19

*Теоретический вопрос:*

Использование метода наименьших квадратов для расчета коэффициентов множественной регрессии.

*Практическое задание:*

В таблице приведены данные (млрд. долл.) по реальному ВНП, реальному объему потребления П и объему инвестиций И для некоторой страны:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ВНП** | 240 | 248 | 261 | 274 | 273 | 269 | 283 | 296 | 312 | 319 |
| **П** | 149 | 154 | 162 | 169 | 167 | 171 | 180 | 188 | 196 | 200 |
| **И** | 38,2 | 41,9 | 46,5 | 52,1 | 48,1 | 38,3 | 45,4 | 52,1 | 56,8 | 57,5 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ВНП** | 318 | 325 | 317 | 327 | 350 | 361 | 372 | 385 | 402 | 412 |
| **П** | 200 | 202 | 205 | 215 | 225 | 235 | 245 | 252 | 261 | 266 |
| **И** | 50,9 | 54,5 | 44,7 | 50,4 | 65,8 | 63,7 | 64 | 76,4 | 71,6 | 71,8 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *И* и *ВНП*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *И = b0 + b1 · ВНП* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Постройте с помощью метода наименьших квадратов уравнение множественной линейной регрессии *И = b0 + b1 ВНП + b2 П*;
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,05;
* проанализируйте, имеются ли признаки наличия мультиколлинеарности.
1. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 95%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения инвестиций при значениях ВНП 300 млрд. долл. и 400 млрд. долл.

## Вариант 20

*Теоретический вопрос:*

Мультиколлинеарность (суть, последствия, методы обнаружения и устранения).

*Практическое задание:*

Имеются данные о деятельности крупнейших компаний США – чистом доходе
(*Y*, млрд. долл.), обороте капитала (*Х1*, млрд. долл.) и численности служащих (*Х2*, тыс. чел.):

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 0,9 | 1,7 | 0,7 | 1,7 | 2,6 | 1,3 | 4,1 | 1,6 |
| **Х1** | 31,3 | 13,4 | 4,5 | 10 | 20 | 15 | 137,1 | 17,9 |
| **Х2** | 43 | 64,7 | 24 | 50,2 | 106 | 96,6 | 347 | 85,6 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 6,9 | 0,4 | 1,3 | 1,9 | 1,9 | 1,4 | 0,4 | 0,8 |
| **Х1** | 165,4 | 2 | 6,8 | 27,1 | 13,4 | 9,8 | 19,5 | 6,8 |
| **Х2** | 745 | 4,1 | 56,8 | 42,7 | 61,8 | 212 | 105 | 33,5 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 1,8 | 0,9 | 1,1 | 1,9 | -0,9 | 1,3 | 2 | 0,6 | 0,7 |
| **Х1** | 27 | 12,4 | 17,7 | 12,7 | 21,4 | 13,5 | 13,4 | 4,2 | 15,5 |
| **Х2** | 142 | 96 | 140 | 59,3 | 131 | 70,7 | 65,4 | 23,1 | 80,8 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х1*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X1* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Постройте с помощью метода наименьших квадратов уравнение множественной линейной регрессии *Ŷ = b0 + b1 Х1 + b2 Х2*;
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* проанализируйте, имеются ли признаки наличия мультиколлинеарности.
1. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 95%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения чистого дохода при значениях оборота капитала 30 млрд. долл. и 100 млрд. долл.

## Вариант 21

*Теоретический вопрос:*

Предпосылки метода наименьших квадратов (условия Гаусса-Маркова) для множественной регрессии.

*Практическое задание:*

В таблице представлены данные о рынке строящегося жилья в Санкт-Петербурге – о цене строящегося жилья (*Y*, млн. руб.), общей площади квартиры (*Х1*, м2) и жилой площади квартиры (*Х2*, м2):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 4,77 | 8,1 | 4,05 | 4,53 | 6,33 | 8,61 | 8,16 | 8,49 | 15,69 | 6,6 |
| **Х1** | 39 | 68,4 | 34,8 | 39 | 54,7 | 74,7 | 71,7 | 74,5 | 137,7 | 40 |
| **Х2** | 20 | 40,5 | 16 | 20 | 28 | 46,3 | 45,9 | 47,5 | 87,2 | 17,7 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 8,4 | 13,5 | 15,3 | 10,32 | 7,41 | 9,24 | 4,77 | 8,7 | 4,62 | 8,58 |
| **Х1** | 53 | 86 | 98 | 62,6 | 45,3 | 56,4 | 37 | 67,5 | 37 | 69 |
| **Х2** | 31,1 | 48,7 | 65,8 | 31,4 | 20,6 | 29,7 | 17,8 | 43,5 | 17,8 | 42,4 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х1*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X1* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Постройте с помощью метода наименьших квадратов уравнение множественной линейной регрессии *Ŷ = b0 + b1 Х1 + b2 Х2*;
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* проанализируйте, имеются ли признаки наличия мультиколлинеарности.
1. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения цены жилья при значениях общей площади квартиры 50 м2 и 90 м2.

## Вариант 22

*Теоретический вопрос:*

Оценка моделей с временными лагами в независимых переменных.

*Практическое задание:*

Имеются данные о продаже однокомнатных квартир на вторичном рынке жилья в Санкт-Петербурге – о цене квартиры (*Y*, млн. руб.), общей площади квартиры (*Х1*, м2) и жилой площади квартиры (*Х2*, м2):

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 3,9 | 4,95 | 5,1 | 4,5 | 4,26 | 3,15 | 6,9 | 3,6 | 4,68 | 3,75 | 3,39 |
| **Х1** | 37 | 60 | 60 | 53 | 35 | 30,3 | 43 | 30 | 35 | 32 | 31 |
| **Х2** | 21,5 | 27 | 30 | 26,2 | 19 | 17,5 | 25,5 | 17,8 | 18 | 17 | 18 |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 3,9 | 6,3 | 3,6 | 3,3 | 3,3 | 6,75 | 7,8 | 5,55 | 3,96 | 7,74 | 5,1 |
| **Х1** | 33 | 53 | 32,2 | 31 | 36 | 48 | 55,5 | 48 | 44,1 | 80 | 60 |
| **Х2** | 19,6 | 26 | 18 | 17,3 | 19 | 29 | 35 | 28 | 30 | 51 | 38 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х2*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X2* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Постройте с помощью метода наименьших квадратов уравнение множественной линейной регрессии *Ŷ = b0 + b1 Х1 + b2 Х2*;
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* проанализируйте, имеются ли признаки наличия мультиколлинеарности.
1. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения цены жилья при значениях общей площади квартиры 35 м2 и 70 м2.

## Вариант 23

*Теоретический вопрос:*

Авторегрессионные модели.

*Практическое задание:*

Изучается влияние стоимости основных (*Х1*, млн. руб.) и оборотных (*Х2*, млн. руб.) средств на величину валового дохода (*Y*, млн. руб.) торговых предприятий:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 203 | 63 | 45 | 113 | 81 | 88 | 110 | 146 | 120 | 237 | 160 | 75 |
| **X1** | 138 | 28 | 17 | 50 | 56 | 102 | 116 | 124 | 114 | 154 | 115 | 98 |
| **X2** | 105 | 56 | 54 | 63 | 52 | 85 | 84 | 82 | 76 | 106 | 88 | 66 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х2*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X2* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Постройте с помощью метода наименьших квадратов уравнение множественной линейной регрессии *Ŷ = b0 + b1 Х1 + b2 Х2*;
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проанализируйте, имеются ли признаки наличия мультиколлинеарности.
1. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 95%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения валового дохода при значениях основных средств 70 млн. руб. и 100 млн. руб.

## Вариант 24

*Теоретический вопрос:*

Оценивание параметров системы эконометрических уравнений.

*Практическое задание:*

Имеются данные о деятельности крупнейших компаний США – о чистом доходе
(*Y*, млрд. долл.), обороте капитала (*Х1*, млрд. долл.) и использованном капитале (*Х2*, млрд. долл.):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 6,6 | 3 | 6,5 | 3,3 | 0,1 | 3,6 | 1,5 | 5,5 | 2,4 | 3 |
| **X1** | 116,9 | 18 | 107,9 | 16,7 | 19,6 | 16,2 | 5,9 | 53,1 | 18,8 | 35,3 |
| **X2** | 63,6 | 6,5 | 50,4 | 15,4 | 9,6 | 13,3 | 5,9 | 27,1 | 11,2 | 16,4 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 4,2 | 2,7 | 1,6 | 2,4 | 3,3 | 1,8 | 2,4 | 1,6 | 1,4 | 0,9 |
| **X1** | 71,9 | 73,6 | 10 | 31,5 | 36,7 | 13,8 | 64,8 | 30,4 | 12,1 | 31,3 |
| **X2** | 32,5 | 25,4 | 6,4 | 12,5 | 14,3 | 6,5 | 22,7 | 15,8 | 9,3 | 18,9 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х1*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X1* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Постройте с помощью метода наименьших квадратов уравнение множественной линейной регрессии *Ŷ = b0 + b1 Х1 + b2 Х2*;
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1;
* проанализируйте, имеются ли признаки наличия мультиколлинеарности.
1. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечные прогнозы среднего значения чистого дохода при значениях использованного капитала 15 млрд. долл. и 50 млрд. долл.

## Вариант 25

*Теоретический вопрос:*

Эконометрический метод. Порядок проведения прикладных исследований в эконометрике.

*Практическое задание:*

В таблице представлены данные о рынке строящегося жилья в Санкт-Петербурге – о цене строящегося жилья (*Y*, млн. руб.), общей площади квартиры (*Х1*, м2) и площади кухни (*Х2*, м2):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 4,77 | 8,1 | 4,05 | 4,53 | 6,33 | 8,61 | 8,16 | 8,49 | 15,69 | 6,6 |
| **Х1** | 39 | 68,4 | 34,8 | 39 | 54,7 | 74,7 | 71,7 | 74,5 | 137,7 | 40 |
| **Х2** | 8,2 | 10,7 | 10,7 | 8,5 | 10,7 | 10,7 | 10,7 | 10,4 | 14,6 | 11 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 8,4 | 13,5 | 15,3 | 10,32 | 7,41 | 9,24 | 4,77 | 8,7 | 4,62 | 8,58 |
| **Х1** | 53 | 86 | 98 | 62,6 | 45,3 | 56,4 | 37 | 67,5 | 37 | 69 |
| **Х2** | 10 | 14 | 13 | 11 | 10,4 | 9,4 | 8,3 | 8,3 | 8,3 | 8,3 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х1*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X1* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,01;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Постройте с помощью метода наименьших квадратов уравнение множественной линейной регрессии *Ŷ = b0 + b1 Х1 + b2 Х2*;
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,01;
* проанализируйте, имеются ли признаки наличия мультиколлинеарности.
1. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечный прогноз среднего значения цены жилья при значениях общей площади квартиры 90 м2 и площади кухни 14 м2.

## Вариант 26

*Теоретический вопрос:*

ANOVA-модели. ANCOVA-модели.

*Практическое задание:*

Имеются данные о деятельности крупнейших компаний США – чистом доходе
(*Y*, млрд. долл.), численности служащих (*Х1*, тыс. чел.) и использованном капитале
(*Х2*, млрд. долл.):

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 0,9 | 1,7 | 0,7 | 1,7 | 2,6 | 1,3 | 4,1 | 1,6 |
| **Х1** | 43 | 64,7 | 24 | 50,2 | 106 | 96,6 | 347 | 85,6 |
| **Х2** | 18,9 | 13,7 | 18,5 | 4,8 | 21,8 | 5,8 | 99 | 20,1 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 6,9 | 0,4 | 1,3 | 1,9 | 1,9 | 1,4 | 0,4 | 0,8 |
| **Х1** | 745 | 4,1 | 56,8 | 42,7 | 61,8 | 212 | 105 | 33,5 |
| **Х2** | 60,6 | 1,4 | 8 | 18,9 | 13,2 | 12,6 | 12,2 | 3,2 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 1,8 | 0,9 | 1,1 | 1,9 | -0,9 | 1,3 | 2 | 0,6 | 0,7 |
| **Х1** | 142 | 96 | 140 | 59,3 | 131 | 70,7 | 65,4 | 23,1 | 80,8 |
| **Х2** | 13 | 6,9 | 15 | 11,9 | 1,6 | 8,6 | 11,5 | 1,9 | 5,8 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х1*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X1* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Постройте с помощью метода наименьших квадратов уравнение множественной линейной регрессии *Ŷ = b0 + b1 Х1 + b2 Х2*;
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,01;
* проанализируйте, имеются ли признаки наличия мультиколлинеарности.
1. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 99%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечный прогноз среднего значения чистого дохода при значениях численности служащих 100 тыс. чел. и стоимости использованного капитала 30 млрд. долл.

## Вариант 27

*Теоретический вопрос:*

Использование фиктивных переменных для описания качественных переменных, имеющих более двух альтернатив.

*Практическое задание:*

По данным 25 стран изучается зависимость индекса человеческого развития (*Y*) от суточной калорийности питания населения (*Х1*, ккал на душу населения) и ожидаемой продолжительности жизни при рождении (*Х2*, число лет):

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 0,904 | 0,922 | 0,763 | 0,923 | 0,918 | 0,906 | 0,905 | 0,545 |
| **Х1** | 3343 | 3001 | 3101 | 3543 | 3237 | 3330 | 3808 | 2415 |
| **Х2** | 77 | 78,2 | 68 | 77,2 | 77,2 | 77,2 | 75,7 | 62,6 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 0,894 | 0,9 | 0,932 | 0,74 | 0,701 | 0,744 | 0,921 | 0,927 |
| **Х1** | 3295 | 3504 | 3056 | 3007 | 2844 | 2861 | 3259 | 3350 |
| **Х2** | 78 | 78,2 | 79 | 67,6 | 69,8 | 68,4 | 77,9 | 78,1 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 0,802 | 0,747 | 0,927 | 0,721 | 0,913 | 0,918 | 0,833 | 0,914 | 0,923 |
| **Х1** | 3344 | 2704 | 3642 | 2753 | 2916 | 3551 | 3177 | 3280 | 3160 |
| **Х2** | 72,5 | 66,6 | 76,7 | 68,8 | 76,8 | 78,1 | 73,9 | 78,6 | 78,5 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х1*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X1* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Постройте с помощью метода наименьших квадратов уравнение множественной линейной регрессии *Ŷ = b0 + b1 Х1 + b2 Х2*;
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* проанализируйте, имеются ли признаки наличия мультиколлинеарности.
1. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 99%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечный прогноз среднего значения индекса человеческого развития при значениях суточной калорийности питания 3050 ккал на душу населения и ожидаемой продолжительности жизни при рождении 75 лет.

## Вариант 28

*Теоретический вопрос:*

Модели множественной нелинейной регрессии.

*Практическое задание:*

Имеются данные о деятельности крупнейших компаний США – о чистом доходе
(*Y*, млрд. долл.), численности служащих (*Х1*, тыс. чел.) и использованном капитале (*Х2*, млрд. долл.):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 6,6 | 3 | 6,5 | 3,3 | 0,1 | 3,6 | 1,5 | 5,5 | 2,4 | 3 |
| **X1** | 222 | 122 | 182 | 125,2 | 29,3 | 141,6 | 17,8 | 151 | 82,3 | 113 |
| **X2** | 63,6 | 6,5 | 50,4 | 15,4 | 9,6 | 13,3 | 5,9 | 27,1 | 11,2 | 16,4 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 4,2 | 2,7 | 1,6 | 2,4 | 3,3 | 1,8 | 2,4 | 1,6 | 1,4 | 0,9 |
| **X1** | 105,4 | 75 | 43,8 | 122,3 | 105 | 53,1 | 50,4 | 80 | 81 | 43 |
| **X2** | 32,5 | 25,4 | 6,4 | 12,5 | 14,3 | 6,5 | 22,7 | 15,8 | 9,3 | 18,9 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х1*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X1* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие автокорреляции остатков графическим методом и методом рядов при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Постройте с помощью метода наименьших квадратов уравнение множественной линейной регрессии *Ŷ = b0 + b1 Х1 + b2 Х2*;
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,01;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста Голдфельда-Квандта при уровне значимости 0,1;
* проанализируйте, имеются ли признаки наличия мультиколлинеарности.
1. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечный прогноз среднего значения чистого дохода при значениях численности служащих 50 тыс. чел. и стоимости использованного капитала 20 млрд. долл.

## Вариант 29

*Теоретический вопрос:*

Причины наличия в регрессионной модели случайного отклонения.

*Практическое задание:*

По данным 25 стран изучается зависимость индекса человеческого развития (*Y*) от ожидаемой продолжительности жизни при рождении (*Х1*, число лет) и ВВП (*Х2*, % к базисному году):

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 0,904 | 0,922 | 0,763 | 0,923 | 0,918 | 0,906 | 0,905 | 0,545 |
| **Х1** | 77 | 78,2 | 68 | 77,2 | 77,2 | 77,2 | 75,7 | 62,6 |
| **Х2** | 15,2 | 15,8 | 20,7 | 17,8 | 15,9 | 14,4 | 14,6 | 28,2 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 0,894 | 0,9 | 0,932 | 0,74 | 0,701 | 0,744 | 0,921 | 0,927 |
| **Х1** | 78 | 78,2 | 79 | 67,6 | 69,8 | 68,4 | 77,9 | 78,1 |
| **Х2** | 17,7 | 15,5 | 15,7 | 18,5 | 22,4 | 23 | 14,2 | 15,2 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Y** | 0,802 | 0,747 | 0,927 | 0,721 | 0,913 | 0,918 | 0,833 | 0,914 | 0,923 |
| **Х1** | 72,5 | 66,6 | 76,7 | 68,8 | 76,8 | 78,1 | 73,9 | 78,6 | 78,5 |
| **Х2** | 18,4 | 22,7 | 18,1 | 20,1 | 17,3 | 16,8 | 19,9 | 16,3 | 14,1 |

Задание:

1. Исследуйте корреляционную зависимость между переменными *У* и *Х1*:
* рассчитайте коэффициент корреляции между переменными, сделайте вывод о силе и направлении связи между ними;
* проверьте сделанные выводы с помощью поля корреляции;
* предложите одну или несколько математических функций, наиболее соответствующих зависимости между переменными.
1. Постройте с помощью метода наименьших квадратов уравнение парной линейной регрессии *Ŷ = b0 + b1 · X1* для описания зависимости между переменными. Приведите экономическую интерпретацию параметров уравнения.
2. Проверьте качество построенной регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие автокорреляции остатков графическим методом и с помощью критерия Дарбина-Уотсона при уровне значимости 0,01;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* сделайте вывод, можно ли использовать линейную модель для прогнозирования. Совпадают ли ваши выводы с предположениями, сделанными в п. 1?
1. Постройте с помощью метода наименьших квадратов уравнение множественной линейной регрессии *Ŷ = b0 + b1 Х1 + b2 Х2*;
2. Проверьте качество новой регрессионной модели:
* проверьте адекватность уравнения регрессии эмпирическим данным;
* проверьте адекватность уравнения регрессии данным генеральной совокупности при уровне значимости 0,05;
* проверьте статистическую значимость коэффициентов уравнения при уровне значимости 0,05;
* проверьте наличие гетероскедастичности графическим методом и с помощью теста ранговой корреляции при уровне значимости 0,01;
* проанализируйте, имеются ли признаки наличия мультиколлинеарности.
1. Сравните качество построенных моделей. Какая из моделей, на ваш взгляд, предпочтительнее для выражения исследуемой зависимости и почему?
2. Рассчитайте 90%-ные доверительные интервалы для теоретических коэффициентов наилучшей регрессии. Сделайте выводы.
3. По наилучшей регрессионной модели рассчитайте точечный прогноз среднего значения индекса человеческого развития при значениях ожидаемой продолжительности жизни при рождении 78 лет и ВВП 20%.

# Рекомендуемая литература

**Основная литература**

1. Бородич С.А. Эконометрика: Учеб. пособие / С.А. Бородич. – Мн.: Новое знание, 2001. – 408 с. – (Экономическое образование)
2. Елисеева И.И. и др. Эконометрика: Учебник. М.: Финансы и статистика, 2004. – 344 с.
3. Практикум по эконометрике: Учебное пособие / ред. И.И. Елисеева. – М.: Финансы и статистика, 2003. – 192 с.
4. В.П. Носко. Эконометрика (книга 1-ая). – М.: Дело, 2011. – 672 с.
5. В.П. Носко. Эконометрика (книга 2-ая). – М.: Дело, 2011. – 576 с.
6. Н.Ш. Кремер, Б.А. Путько. Эконометрика: Учебник для вузов. – М.: ЮНИТИ-ДАНА, 2002. – 311 с.

**Дополнительная литература**

1. Орлов А.И. Эконометрика: Учебник для вузов. – М.: Изд-во «Экзамен», 2003. – 567 с.
2. Катышев П.К., Пересецкий А.А. Сборник задач к начальному курсу эконометрики. – М.: Дело, 1999.
3. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. Учебное пособие. М.: Дело, 1998.
4. Катышев П.К. Сборник задач к начальному курсу эконометрики. – М.: Дело, 1999. – 72 с.
5. Кулинич Е.И. Эконометрия. – М.: Финансы и статистика, 2001. – 304 с.
6. Адамов В.Е. и др. Экономика и статистика фирм. – М: Финансы и статистика, 2001.

# Приложение А

**Пример оформления титульного листа**

Министерство образования и науки Российской Федерации

ФГБОУ ВПО «Российский химико –технологический университет им. Д.И. Менделеева»

Новомосковский институт (филиал)

кафедра «Экономика, финансы и бухгалтерский учет»

**Курсовая работа**

**по курсу «Эконометрика»**

**на тему «Построение регрессионной модели**

**экономического объекта»**

**Вариант \_\_\_**

Студент

Шифр

Преподаватель

Новомосковск 20\_\_

# Приложение Б

**Информационные технологии Microsoft Excel**

**Функция СУММ**

***Синтаксис:***

СУММ (число1; число2; …)

или

СУММ (диапазон чисел)

***Результат:***

Рассчитывает сумму заданных чисел.

***Аргументы:***

* *число1, число2, …:* суммируемые числа.
* *диапазон\_чисел:* диапазон ячеек, который содержит суммируемые числа.

**Функция СРЗНАЧ**

***Синтаксис:***

СРЗНАЧ (диапазон\_ячеек)

***Результат:***

Рассчитывает выборочное среднее для набора данных.

***Аргументы:***

* *диапазон\_ячеек:* диапазон ячеек, в котором содержится набор данных.

**Функция КОРЕНЬ**

***Синтаксис:***

КОРЕНЬ (число)

***Результат:***

Рассчитывает квадратный корень из заданного числа.

***Аргументы:***

* *число:* число, корень из которого рассчитывается.

# Приложение В

**ИСПОЛЬЗОВАНИЕ ФУНКЦИИ «ЛИНЕЙН» ДЛЯ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ УРАВНЕНИЯ РЕГРЕССИИ**

***Определение параметров парной регрессии***

***Синтаксис:***

ЛИНЕЙН (известные\_значения\_у; известные\_значения\_х; константа; статистика)

***Результат:***

На основе метода наименьших квадратов рассчитывается массив данных, описывающих уравнение линейной регрессии.

***Аргументы:***

- *известные\_значения\_у*: диапазон значений (в столбце или в строке) результирующего показателя Y;

- *известные\_значения\_х*: диапазон значений (в столбце или в строке) фактора Х;

- *константа*: логическое значение: если оно равно 0, свободный член b0 равен 0;

если оно равно 1, то b0 вычисляется обычным образом.

- *статистика*: логическое значение:

если оно равно 0, то функция рассчитывает только коэффициенты b0 и b1;

если оно равно 1, то функция рассчитывает дополнительную регрессионную статистику.

В случае парной регрессии возвращаемый массив данных будет иметь вид, указанный в таблице В.1:

Таблица В.1

|  |  |
| --- | --- |
| *b1* | *b0* |
| *Sb1* | *Sb0* |
| *R2* | *Se* |
| *F* | *ν* |
| *ssоб* | *ssост* |

b1 и b0 – значения коэффициентов регрессии;

Sb1 и Sb0 – стандартные ошибки коэффициентов регрессии;

R2 – величина коэффициента детерминации;

Se – величина стандартной ошибки регрессии;

F – значение критерия Фишера для проверки гипотезы о значимости R2;

*ν -* число степеней свободы, равное n-2;

*ssоб*- объясненная сумма квадратов отклонений – Σ(ŷi – ‾y)2;

*ssост* – остаточная сумма квадратов отклонений - Σеi2.

Чтобы функция регрессии возвратила массив из десяти ячеек, необходимо ввести ее как *функцию массива*. Это можно осуществить двумя способами:

**1.** Выделяются десять ячеек, в которые нужно поместить результат функции (как в таблице В.1 – 2 столбца и 5 строк). Затем вводится функция ЛИНЕЙН (сразу во все выделенные ячейки), причем закончить ввод нужно комбинацией клавиш Ctrl + Shift + Enter.

**2.** Функция ЛИНЕЙН вводится в одну ячейку – эта ячейка будет верхней левой ячейкой массива результатов. Затем выделяются десять ячеек, в которые нужно поместить результат, начиная с ячейки, где введена функция (как в таблице В.1 – 2 столбца и 5 строк). Выделив ячейки, нужно нажать клавишу F2, а затем – комбинацию клавиш Ctrl + Shift + Enter.

***Определение параметров множественной регрессии***

**Функция ЛИНЕЙН**

***Синтаксис:***

ЛИНЕЙН (известные\_значения\_у; известные\_значения\_х; константа; статистика)

***Результат:***

На основе метода наименьших квадратов рассчитывается массив данных, описывающих уравнение линейной регрессии.

***Аргументы:***

- *известные\_значения\_у*: диапазон значений (в столбце или в строке) результирующего показателя Y;

- *известные\_значения\_х*: диапазон значений факторов (одновременно выделяются все столбцы или строки, содержащие значения факторов Х1, Х2 и т.д.);

- *константа*: логическое значение: если оно равно 0, свободный член b0 равен 0;

если оно равно 1, то b0 вычисляется обычным образом.

- *статистика*: логическое значение:

если оно равно 0, то функция рассчитывает только коэффициенты регрессии;

если оно равно 1, то функция рассчитывает дополнительную регрессионную статистику.

В случае множественной регрессии возвращаемый массив данных будет иметь вид, указанный в таблице В.2:

Таблица В.2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *bm* | *bm-1* | *…* | *b1* | *b0* |
| *Sbm* | *Sb(m-1)* | *…* | *Sb1* | *Sb0* |
| *R2* | *Se* |  |  |  |
| *F* | *ν* |  |  |  |
| *ssоб* | *ssост* |  |  |  |

b0, b1, …, bm – значения коэффициентов множественной регрессии;

Sb0, Sb1, …, Sbm – стандартные ошибки коэффициентов множественной регрессии;

R2 – величина коэффициента детерминации;

Se – величина стандартной ошибки регрессии;

F – значение критерия Фишера для проверки гипотезы о значимости R2;

*ν -* число степеней свободы, равное n-m-1;

*ssоб*- объясненная сумма квадратов отклонений – Σ(ŷi – ‾y)2;

*ssост* – остаточная сумма квадратов отклонений - Σеi2.

Ячейки, которые не содержат данных, отмечены знаком «#Н/Д» (нет данных).

Чтобы функция регрессии возвратила массив ячеек, необходимо ввести ее как *функцию массива*. Это можно осуществить двумя способами, описанными выше.