

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Волков В.В.

Должность: Ректор

Дата подписания: 18.08.2025 15:58:22

Уникальный программный ключ:

ed68fd4b85b778e0f0b1bfea5dbc56cf4148f1229917e799a70c51517ff6d991

**Автономная некоммерческая образовательная организация высшего образования**

**«Европейский университет в Санкт-Петербурге»**

УТВЕРЖДАЮ:

Ректор

В.В. Волков

Протокол УС №

от 26.02.2025 г.

**Рабочая программа дисциплины  
Основы регрессионного анализа**

образовательная программа  
направление подготовки  
**38.04.01 Экономика**

**направленность (профиль) программы  
«Данные, знания, экономика, нарративы»  
уровень высшего образования – магистратура**

**Программа двух квалификаций:**

- «магистр» по направлению подготовки **38.04.01 Экономика;**
- дополнительная квалификация – «магистр» по направлению подготовки **09.04.03 Прикладная информатика**

язык обучения – русский  
форма обучения - очная

**Санкт-Петербург**

**Автор:**

Соколов М.В., кандидат экономических наук, старший научный сотрудник Школы вычислительных социальных наук АНООВО «ЕУСПб»

**Рецензент:** Вымятнина Ю.В., кандидат экономических наук, профессор Школы вычислительных социальных наук АНООВО «ЕУСПб»

Рабочая программа дисциплины **«Основы регрессионного анализа»**, входящей в образовательную программу уровня магистратуры «Данные, знания, экономика, нарративы», утверждена на заседании Совета Школы вычислительных социальных наук.

Протокол заседания № 4 от 25.02.2025 года.

## **АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ**

Дисциплина «**Основы регрессионного анализа**» является дисциплиной базовой части Блока 1 «Дисциплины (модули)» образовательной программы «Данные, знания, экономика, нарративы».

Дисциплина «**Основы регрессионного анализа**» посвящена изучению базовых тем регрессионного анализа. Дисциплина дает студентам представление о содержательном смысле линейных регрессионных моделей, их основных допущениях, инструментах диагностики этих допущений, оценивании моделей, а также интерпретации полученных результатов.

Общая трудоемкость освоения дисциплины составляет 2 зачетных единицы, 72 часа.

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....	5
2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....	5
3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ .....	6
4. ОБЪЕМ ДИСЦИПЛИНЫ .....	6
5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ .....	7
5.1 Содержание дисциплины.....	7
5.2 Структура дисциплины.....	8
6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ.....	8
6.1 Общие положения .....	8
6.2 Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины .....	8
6.3 Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине.....	9
6.4 Перечень литературы для самостоятельной работы обучающегося:.....	9
6.5 Перечень учебно-методического обеспечения для самостоятельной работы.....	10
7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ.....	10
7.1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации .....	10
7.2 Контрольные задания для текущей аттестации.....	12
7.3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации.....	15
7.4 Типовые задания к промежуточной аттестации.....	17
7.5 Средства оценки индикаторов достижения компетенций.....	24
8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА .....	25
8.1. Основная литература.....	25
8.2. Дополнительная литература.....	25
9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА .....	25
9.1 Программное обеспечение .....	25
9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины:.....	25
9.3 Лицензионные электронные ресурсы библиотеки Университета .....	26
9.4 Электронная информационно-образовательная среда Университета.....	26
10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА .....	27
ПРИЛОЖЕНИЕ 1 .....	28

## **1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ**

**Целью** освоения дисциплины «Основы регрессионного анализа» является знакомство студентов с основные понятиями и инструментами регрессионного анализа и их использованием в фундаментальных и прикладных исследованиях.

**Задачи** освоения дисциплины «Основы регрессионного анализа» включают:

1. изучение основных инструментов регрессионного анализа,
2. развитие навыков самостоятельного построения эконометрических моделей,
3. развитие навыков проверки адекватности этих моделей,
4. развитие навыков экономической интерпретации полученных результатов.

## **2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ**

В результате изучения учебной дисциплины обучающийся должен овладеть следующими компетенциями: общепрофессиональными (ОПК). Планируемые результаты формирования компетенций и индикаторы их достижения в результате освоения дисциплины представлены в Таблице 1.

Таблица 1

**Планируемые результаты освоения дисциплины, соотнесенные с индикаторами достижения компетенций обучающихся**

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
ОПК-2-Э Способен применять продвинутые инструментальные методы экономического анализа в прикладных и (или) фундаментальных исследованиях	<p>ИД.ОПК-2.1. – Владеет современными методами экономического анализа, математической статистики и эконометрики для решения теоретических и прикладных задач</p> <p>ИД.ОПК-2.2 - Работает с национальными и международными базами данных с целью поиска необходимой информации об экономических явлениях и процессах</p> <p>ИД.ОПК-2.3 - Обрабатывает статистическую информацию и получает статистически обоснованные выводы</p>	<p>Знать: современные методы экономического анализа, математической статистики и эконометрики для решения теоретических и прикладных задач в профессиональной области; национальные и международные базы данных об экономических явлениях и процессах 3 (ОПК-2)</p> <p>Уметь: решать теоретические и прикладные задачи профессиональной области, применяя методы экономического анализа, математической статистики и эконометрики У (ОПК-2)</p> <p>Владеть: навыками обработки и анализа статистической информации, а также представления полученных результатов статистического и экономического анализа в соответствии с установленными требованиями В (ОПК-2)</p>
ОПК-5-Э Способен использовать современные информационные технологии и программные средства при решении профессиональных задач	<p>ИД.ОПК-5.1. Обосновывает выбор информационно-коммуникационных технологий для постановки и решения задач экономического исследования</p> <p>ИД.ОПК-5.2. Предлагает постановку задач по поиску, обобщению экономических данных</p> <p>ИД.ОПК-5.3. Осуществляет обоснованный выбор методов и в необходимых случаях – программных средств – для</p>	<p>Знать: перспективные информационно-коммуникационные технологии и программные средства обработки и представления информации, необходимые для комплексной постановки и высокоэффективного решения задач профессиональной деятельности 3 (ОПК-5)</p> <p>Уметь: осуществлять поиск верифицированных источников информации по выбранной области исследований и обработку полученных</p>

<b>Код и наименование компетенции</b>	<b>Индикаторы достижения компетенции</b>	<b>Результаты обучения (знать, уметь, владеть)</b>
	<p>обработки экономической информации, проверки гипотез исследования и оценки надёжности полученных данных ИД.ОПК-5.4. Создаёт и поддерживает нормативно-методическую и информационную базу исследований по диагностике, оценке, оптимизации социально-экономических показателей, процессов и отношений ИД.ОПК-5.5. Устанавливает правила, регламентирующие порядок и условия доступа к социально-экономической информации, и контролирует их выполнение</p>	<p>данных путем применения актуального программного обеспечения для эффективного достижения поставленных профессиональных целей У (ОПК-5)</p> <p>Владеть: навыком применения современных информационно-коммуникационных технологий и программных средств для решения профессиональных задач В (ОПК-5)</p>

В результате освоения дисциплины магистрант должен:

**знатъ:**

- основные понятия, связанные с регрессионным анализом,
- основные методы оценивания и диагностики моделей регрессионного анализа;
- основные статистические критерии, используемые в регрессионном анализе;

**уметь:**

- оценивать модели линейной регрессии (с использованием статистических пакетов);
- проводить диагностику регрессионной модели;
- интерпретировать результаты оценки модели.

**владеть:**

- навыками использования эконометрических пакетов для оценки регрессионных моделей и их диагностики.

### **3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ**

Дисциплина «Основы регрессионного анализа» является дисциплиной обязательной части Блока 1 «Дисциплины (модули)» образовательной программы «Данные, знания, экономика, нарративы». Курс читается во втором модуле, форма промежуточной аттестации – зачет с оценкой.

Для успешного освоения данной дисциплины требуются знания, полученные в рамках изучения дисциплин бакалавриата/специалитета по направлению подготовки.

Знания, умения и навыки, полученные при освоении данной дисциплины, применяются магистрантами в процессе прохождения учебной и производственной практики, выполнения выпускной квалификационной работы.

### **4. ОБЪЕМ ДИСЦИПЛИНЫ**

Общая трудоемкость освоения дисциплины составляет 2 зачетных единиц, 72 часа.

Таблица 2

**Объем дисциплины**

Типы учебных занятий и самостоятельная работа	Объем дисциплины										
	Всего	Модуль									
	1	2	3	4	5	6	7	8	9	10	
<b>Контактная работа обучающихся с преподавателем в соответствии с УП:</b>	<b>24</b>	-	24	-	-	-	-	-	-	-	-
Лекции (Л)	<b>12</b>	-	12	-	-	-	-	-	-	-	-
Практические занятия (ПЗ)	<b>12</b>	-	12	-	-	-	-	-	-	-	-
<b>Самостоятельная работа (СР)</b>	<b>48</b>	-	48	-	-	-	-	-	-	-	-
<b>Промежуточная аттестация</b>	<b>форма</b>	<b>Зачет с оценкой</b>	<b>Зачет с оценкой</b>	-	-	-	-	-	-	-	-
	<b>час.</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>
<b>Общая трудоемкость дисциплины (час./з.е.)</b>	<b>72/2</b>	-	<b>72/2</b>	-	-	-	-	-	-	-	-

**5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ**

Содержание дисциплины соотносится с планируемыми результатами обучения по дисциплине: через задачи, формируемые компетенции и их компоненты (знания, умения, навыки – далее ЗУВ) по средствам индикаторов достижения компетенций в соответствии с Таблицей 3.

**5.1 Содержание дисциплины**

Таблица 3

**Содержание дисциплины**

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)
1	Множественная регрессия – основные понятия	Этапы эконометрического исследования и построения эконометрической модели. Типы данных и типы моделей.  Линейная регрессионная модель. Метод наименьших квадратов. Коэффициент детерминации и скорректированный детерминации. Теорема Гаусса-Маркова. Проверка гипотез о коэффициентах. Доверительный интервал. Интерпретация коэффициентов. Прогнозирование.  Фиктивные переменные и их использование. Тесты Чоу на стабильность параметров. Ошибки спецификации (пропущенные переменные, включение лишних переменных, выбор формы модели) и их последствия, RESET тест Рамсея.	ОПК-2-Э ОПК-5-Э	ИД.ОПК-2.1. ИД.ОПК-2.2 ИД.ОПК-2.3 ИД.ОПК-5.1. ИД.ОПК-5.3. ИД.ОПК-5.4. ИД.ОПК-5.5.	3 (ОПК-2) У (ОПК-2) В (ОПК-2) 3 (ОПК-5) У (ОПК-5) В (ОПК-5)
2	Ослабление предпосылок классической линейной регрессионной модели.	Гетероскедастичность: тестирование, оценивание модели в условиях гетероскедастичности.  Автокорреляция: тестирование, оценивание модели в условиях автокорреляции.	ОПК-2-Э ОПК-5-Э	ИД.ОПК-2.1. ИД.ОПК-2.2 ИД.ОПК-2.3 ИД.ОПК-5.1. ИД.ОПК-5.3. ИД.ОПК-5.4. ИД.ОПК-5.5.	3 (ОПК-2) У (ОПК-2) В (ОПК-2) 3 (ОПК-5) У (ОПК-5) В (ОПК-5)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)
		Стохастические регрессоры. Эндогенность. Метод инструментальных переменных. Тест Хаусмана.			

## 5.2 Структура дисциплины

Таблица 4

### Структура дисциплины

№ п/п	Наименование тем (разделов)	Объем дисциплины, час.				Форма текущего контроля успеваемости*, промежуточной аттестации	
		Всего	Контактная работа обучающихся с преподавателем по типам учебных занятий в соответствии с УП	СР			
		Л	ПЗ	СРП			
<b>Очная форма обучения</b>							
Тема 1	Множественная регрессия – основные понятия	52	8	8	-	36	ДЗ
Тема 2	Ослабление предпосылок классической линейной регрессионной модели.	20	4	4	-	12	ДЗ
<b>Промежуточная аттестация</b>		-	-	-	-	-	Зачет с оценкой
<b>Всего:</b>		<b>72/2</b>	<b>12</b>	<b>12</b>	<b>-</b>	<b>48</b>	<b>-</b>

\*Примечание: формы текущего контроля успеваемости: домашнее задание (ДЗ).

## 6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

### 6.1 Общие положения

Знания и навыки, полученные в результате лекций и семинарских занятий, закрепляются и развиваются в результате повторения материала, усвоенного в аудитории, путем чтения текстов и исследовательской литературы (из списков основной и дополнительной литературы) и их анализа.

Самостоятельная работа является важнейшей частью процесса высшего образования. Ее следует осознанно организовать, выделив для этого необходимое время и соответственным образом организовав рабочее пространство. Важнейшим элементом самостоятельной работы является проработка материалов прошедших занятий (анализ конспектов, чтение рекомендованной литературы) и подготовка к следующим лекциям/семинарским занятиям. Литературу, рекомендованную в программе курса, следует, по возможности, читать в течение всего семестра, концентрируясь на обусловленных программой курса темах.

Существенную часть самостоятельной работы магистранта представляет самостоятельное изучение вспомогательных учебно-методических изданий, лекционных конспектов, интернет-ресурсов и пр. Подготовка к семинарским занятиям является важной формой работы магистранта. Самостоятельная работа может вестись как индивидуально, так и при содействии преподавателя.

### 6.2 Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины

**Тема 1. Множественная регрессия – основные понятия.**

1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 18 часа.
2. Подготовка к занятиям по предложенным для обсуждения вопросам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 18 часа. Итого: 36 часов.

## **Тема 2. Ослабление предпосылок классической линейной регрессионной модели.**

1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 6 часов.

2. Подготовка к занятиям по предложенным для обсуждения вопросам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 6 часов. Итого: 12 часов.

## **6.3 Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине**

1. Типы статистических данных и типы моделей.
2. Линейная регрессионная модель. Формулировка задачи и используемые предпосылки.
3. Метод наименьших квадратов.
4. Коэффициент детерминации и скорректированный коэффициент детерминации.
5. Теорема Гаусса-Маркова.
6. Проверка значимости уравнения в целом. Проверка гипотез о коэффициентах.
7. Доверительный интервал для коэффициента.
8. Интерпретация коэффициентов модели.
9. Прогнозирование.
10. Фиктивные переменные и их использование.
11. Тесты Чоу на стабильность параметров.
12. Ошибки спецификации (пропущенные переменные, включение лишних переменных, выбор формы модели) и их последствия. RESET тест Рамсея.
13. Ослабление предпосылок классической линейной регрессионной модели: гетероскедастичность (тестирование гетероскедастичности, оценивание модели в условиях гетероскедастичности).
14. Ослабление предпосылок классической линейной регрессионной модели: автокорреляция (тестирование автокорреляции, оценивание модели в условиях автокорреляции).
15. Стохастические регрессоры. Эндогенность и метод инструментальных переменных. Тест Хаусмана.

## **6.4 Перечень литературы для самостоятельной работы обучающегося:**

1. Я. Р. Магнус, П. К. Катышев, А. А. Пересецкий. Эконометрика. Начальный курс: учебник / Академия народного хозяйства при Правительстве РФ. - 8-е изд. - Москва: Дело, 2007. - 503 с.
2. Грин У.Г. Эконометрический анализ. Книга 1 [Текст] / У.Г. Грин. - М.: ИД «Дело», 2016. - 760 с.
3. Кэмерон, Колин. Микроэконометрика: методы и их применение [Текст]: учебник для вузов: в 2 кн.: перевод с английского / Э.К. Кэмерон, П.К. Триведи; ред. пер. Б. Демешев; Российская академия народного хозяйства и государственной службы при Президенте РФ. - М.: Дело, 2015. - (Академический учебник).

## **6.5 Перечень учебно-методического обеспечения для самостоятельной работы**

Для обеспечения самостоятельной работы магистрантов по дисциплине «Основы регрессионного анализа» разработано учебно-методическое обеспечение в составе:

1. Контрольные задания для подготовки к процедурам текущего контроля (п. 7.2 Рабочей программы).
2. Типовые задания для подготовки к промежуточной аттестации (п. 7.4 Рабочей программы).
3. Рекомендуемые основная, дополнительная литература, Интернет-ресурсы и справочные системы (п. 8, 9 Рабочей программы).
4. Рабочая программа дисциплины размещена в электронной информационно-образовательной среде Университета на электронном учебно-методическом ресурсе АНООВО «ЕУСПб» — образовательном портале LMS Sakai — Sakai@EU.

## **7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ**

### **7.1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации**

Информация о содержании и процедуре текущего контроля успеваемости, методике оценивания знаний, умений и навыков обучающегося в ходе текущего контроля доводятся научно-педагогическими работниками Университета до сведения обучающегося на первом занятии по данной дисциплине.

Текущий контроль предусматривает подготовку магистрантов к каждому практическому занятию, активное слушание на лекциях, выполнение магистрантами домашних заданий. Магистрант должен присутствовать на занятиях, отвечать на поставленные вопросы, показывая, что прочитал разбираемую литературу, представлять содержательные реплики по темам обсуждения.

Текущий контроль проводится в форме выполнения магистрантами домашних заданий, демонстрирующих степень знакомства магистрантов с дополнительной литературой.

Таблица 5  
**Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации**

Наименование разделов	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Тема 1. Множественная регрессия – основные понятия	1. ОПК-2-Э ОПК-5-Э	ИД.ОПК-2.1. ИД.ОПК-2.2 ИД.ОПК-2.3 ИД.ОПК-5.1. ИД.ОПК-5.3. ИД.ОПК-5.4. ИД.ОПК-5.5.	З (ОПК-2) У (ОПК-2) В (ОПК-2) З (ОПК-5) У (ОПК-5) В (ОПК-5)	Домашнее задание	зачтено/ не зачтено
Тема 2. Ослабление предпосылок классической линейной регрессионной модели.	ОПК-2-Э ОПК-5-Э	ИД.ОПК-2.1. ИД.ОПК-2.2 ИД.ОПК-2.3 ИД.ОПК-5.1. ИД.ОПК-5.3. ИД.ОПК-5.4. ИД.ОПК-5.5.	З (ОПК-2) У (ОПК-2) В (ОПК-2) З (ОПК-5) У (ОПК-5) В (ОПК-5)	Домашнее задание	зачтено/ не зачтено

При освоении дисциплины каждая из форм текущего контроля оценивается с использованием 100-балльной шкалы с последующим переводом в бинарную систему для получения результатов текущего контроля, фиксирующих ход образовательного процесса, согласно Таблице 6. Максимально возможное количество баллов для домашней работы – 40 баллов.

Таблица 6  
**Критерии оценивания**

<b>Формы текущего контроля успеваемости</b>	<b>Описание</b>	<b>Показатели</b>	<b>Количество баллов по 100-балльной шкале</b>	<b>Результаты текущего контроля</b>
Домашнее задание.	Домашнее задание представляет собой набор задач по пройденным темам к моменту отчетного занятия. Задачи, предлагаемые студентам, дают им представление о сложности заданий, предлагаемых на контрольной работе. Студенты решают задачи дома и приносят их (или отправляют преподавателю по электронной почте) к поставленному сроку (дедлайну).	Магистрант глубоко и прочно усвоил программный материал; исчерпывающе, последовательно, четко и логически стройно излагает его в письменной домашней работе, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами знаний.	91-100% от максимального возможного числа баллов	
		Магистрант твердо знает материал, грамотно и по существу излагает его в письменной домашней работе, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения.	51-90% от максимального возможного числа баллов	зачтено
		Магистрант имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в ответах письменной домашней работы.	25-50% от максимального возможного числа баллов	
		Магистрант не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями отвечает на вопросы письменной домашней работы.	1-24% от максимального возможного числа баллов	не зачтено

## 7.2 Контрольные задания для текущей аттестации

### Примерный материал домашних заданий

#### Тема 1: Множественная регрессия – основные понятия

1. Найдите оценку метода наименьших квадратов параметра  $b$  в регрессии  $y = b + \varepsilon$ , где  $y$  – зависимая переменная,  $b$  – константа подлежащая оценке,  $\varepsilon$  – ошибки. Чему равен коэффициент детерминации в оцененной модели?

2. Изменение единиц измерения зависимой переменной в регрессионной модели  $y = b_0 + b_1 x_1 + b_2 x_2 + \varepsilon$  НЕ приведёт к изменению (выберите верный вариант ответа)

- оценки коэффициента  $b_0$
- оценки коэффициентов  $b_1$  и  $b_2$
- коэффициента детерминации
- общей суммы квадратов остатков

3. Пусть QCOFFEE – спрос на кофе, PCOFFEE – цена кофе, PTEA – цена чая, PCREAMER – цена сливок, INC – доход потребителей.

**Модель 1:** МНК, использованы наблюдения 1-21

Зависимая переменная: lnQCOFFEE

	Коэффициент	Ст. ошибка	t- статистика	P-значение
const	-2,43973	1,59313	-1,5314	0,14520
lnPCOFFEE	-3,31253	0,688323	-4,8125	???????
lnPTEA	1,5942	0,571989	???????	0,01319
lnPCREAMER	0,0693744	?????????	0,1117	0,91247
lnINC	?????????	0,0993055	3,0110	0,00829
Среднее зав. перемен	1,491580	Ст. откл. зав. перемен		0,504265
Сумма кв. остатков	1,320197	Ст. ошибка модели		0,287250
R-квадрат	0,740408	Испр. R-квадрат		??????????
F(4, 16)	?????????	P-значение (F)		0,000143
Лог.правдоподобие	-0,746924	Крит. Акаике		11,49385
Крит. Шварца	16,71646	Крит. Хеннана-Куинна		12,62729

- 1) Восстановите цифры, заменённые на знак вопроса.
- 2) Выпишите оцененное уравнение.
- 3) Проверьте значимость уравнения в целом.
- 4) Проверьте значимость отдельных коэффициентов.
- 5) Интерпретируйте коэффициенты.
- 6) Является ли спрос на кофе эластичным?
- 7) Является ли чай субститутом кофе?
- 8) Является ли кофе нормальным товаром?
- 9) Является ли кофе предметом роскоши?
- 10) Являются ли сливки комплементарным товаром?

4. В файле tropic.csv содержатся данные для 84 стран мира:

GROWTH – среднегодовой темп экономического роста за 1965–1990 годы,  
LOGGDP65 – логарифм ВНП на душу населения в 1965 году,  
LND100km – доля территории, находящейся не далее 100 км от побережья,  
POP100km – доля населения, проживающего не далее 100 км от побережья,  
TROPICAR – доля территории, находящейся в тропиках,  
TROPOPOP – доля населения, проживающего в тропиках,  
OPEN – доля лет в 1965–1990 гг., в течение которых страна была открыта для международной торговли,  
POLINST – индекс качества политических институтов (от 0 – низкое качество, до 10 – высокое качество).

- a) Оцените зависимость темпа экономического роста от всех остальных переменных с помощью множественной линейной регрессии.
- b) Как можно интерпретировать коэффициенты при объясняющих переменных в этой регрессии?
- c) Какие из коэффициентов статистически незначимы?
- d) Являются ли незначимые коэффициенты незначимыми **совокупно** (в смысле F-теста)?
- e) Можно ли по результатам оценки утверждать, что более богатые страны (по уровню ВНП на душу населения в 1965 году) росли медленнее?
- f) Можно ли утверждать, что экономический рост был ниже в странах с большей долей территории или населения в тропиках?

5. Файл real\_estate\_price\_size\_year\_view.csv содержит данные:

price – цена недвижимости

size – размер недвижимости (в квадратных метрах)

year – год постройки

view – наличие вида на море (текстовая переменная «Sea view» или «No sea view»)

- a) Загрузите данные и создайте дамми-переменную для переменной view, где 1 соответствует «Sea view», а 0 — «No sea view».
- b) Постройте модель множественной линейной регрессии, где зависимая переменная — price, а независимые переменные — size, year и дамми-переменная view.
- c) Интерпретируйте коэффициенты модели. Особое внимание уделите коэффициенту при дамми-переменной view.
- d) Проверьте значимость коэффициентов модели с использованием t-тестов.
- e) Постройте график зависимости цены от размера недвижимости, разделив данные на группы по наличию вида на море
- f) Проверьте, зависит ли эффект размера недвижимости на её цену от наличия вида на море, включив в модель взаимодействие между переменными size и view.

## Тема 2: Ослабление предпосылок классической линейной регрессионной модели.

1. Что из следующего является следствием гетероскедастичности?
  - a) МНК-оценки коэффициентов регрессии смещены.
  - b) МНК-оценки коэффициентов регрессии несостоительны.
  - c) МНК-оценки коэффициентов регрессии неэффективны.
  - d) МНК-оценка матрицы ковариаций коэффициентов регрессии является смещённой и несостоительной.

- e) Применение стандартных тестов (Стьюдента, Фишера) некорректно.
2. Что из следующего является следствием автокорреляции?
- МНК-оценки коэффициентов регрессии смещены.
  - МНК-оценки коэффициентов регрессии несостоительны.
  - МНК-оценки коэффициентов регрессии неэффективны.
  - МНК-оценка матрицы ковариаций коэффициентов регрессии является смещённой и несостоительной.
- e) Применение стандартных тестов (Стьюдента, Фишера) некорректно.

2. Файл `financial_wealth13.csv` содержит данные

`NETTFA` – общие чистые финансовые активы человека, в тысячах денежных единиц

`INC` – доход человека, в тысячах денежных единиц

`age` – возраст, лет

`agesq` – возраст в квадрате

`e401k` = 1, если человек участвует в накопительном пенсионном плане, и = 0, если не участвует.

Оцените уравнение

$$\text{NETTFA} = C(1) * \text{INC} + C(2) * \text{age} + C(3) * \text{agesq} + C(4) * \text{e401k} + \varepsilon$$

- Протестируйте его на наличие гетероскедастичности всеми известными вам способами.
- Если гетероскедастичность будет обнаружена, предложите способы того, как решить эту проблему и реализуйте это.
- Прокомментируйте свои ответы и решения.

3. Используя данные Росстата и Банка России за период с 2008 по 2024 год, проанализируйте зависимость потребления домашних хозяйств от располагаемых доходов и процентной ставки.

Переменные:

`cons` — расходы на конечное потребление домашних хозяйств, млрд руб.

`income` — располагаемые денежные доходы населения, млрд руб.

`rate` — средняя процентная ставка по кредитам физическим лицам

- a) Постройте линейную регрессионную модель:

$$\text{cons} = \beta_0 + \beta_1 * \text{income} + \beta_2 * \text{rate} + \varepsilon$$

- Проверьте наличие автокорреляции первого порядка в остатках модели с помощью теста Дарбина–Уотсона.
- Обсудите, как автокорреляция влияет на:
  - оценку коэффициента;
  - корректность стандартных ошибок;
  - доверительные интервалы и t-статистики.

4. Файл `fertility.csv` содержит следующие данные:

HOURSWM – количество рабочих часов женщины в неделю (зависимая переменная),

MOREKIDS – дамми: 1, если у женщины более двух детей; 0 – иначе

AGEFSTM – возраст матери на момент рождения первого ребёнка

MULTI2ND – дамми: 1, если второй и третий ребёнок — близнецы

BOY1ST Дамми: 1, если первый ребёнок — мальчик

BOY2ND Дамми: 1, если второй ребёнок — мальчик

- a. Оцените МНК-регрессию hourswm на morekids и agefstm.
- b. Интерпретируйте коэффициент при переменной morekids.
- c. Есть ли основания считать, что этот коэффициент смещен и нам нужны инструментальные переменные для morekids?
- d. Постройте переменные two\_girls — оба первых ребёнка девочки, two\_boys — оба первых ребёнка мальчики.
- e. Обсудите, могут ли multi2nd, two\_girls и two\_boys выступать хорошими инструментами для morekids. Учитывайте два ключевых свойства инструментов:
  - релевантность;
  - экзогенность.
- f. Проведите тест на слабость инструментов.
- g. Проведите тест Хаусмана вручную.
- h. Если хотя бы один из инструментов выглядит слабым или подозрительным по смыслу — уберите его и повторите оценку.
- i. Если все инструменты проходят ваши проверки — сохраните исходную спецификацию.
- j. Сравните оценки параметров по МНК и IV-моделям. Как изменилась оценка при morekids? Как вы это интерпретируете?

### **7.3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации**

**Форма промежуточной аттестации – зачет с оценкой в форме тестирования.**

Перед зачетом с оценкой проводятся консультации, на которых преподаватель отвечает на вопросы магистрантов.

В результате промежуточного контроля знаний студенты получают аттестацию по дисциплине.

При выставлении оценки по итогам освоения дисциплины учитываются как результаты тестирования, так и текущего контроля успеваемости. Максимальная оценка по итогам освоения дисциплины составляет 100 баллов и определяется суммой баллов, полученных обучающимся за домашние задания (40 баллов максимально) и за тестирование (60 баллов максимально).

При вычислении оценки обучающегося по итогам освоения дисциплины используется процедура округления до целого.

В рамках тестирования студентам будет необходимо решить 20 заданий базовой, повышенной и высокой сложности, из которых 8 заданий являются заданиями базовой сложности, 6 заданий повышенной сложности и 6 заданий высокой сложности. Задания распределяются между компетенциями в равной пропорции для всех уровней сложности.

В базовых заданиях проверяются знания студента по пройденному курсу, в остальных – способность применять знания и логически обосновать свой выбор.

**Тест оценивается в баллах в соответствии со следующими критериями:**

**Задания закрытого типа**

*Базовый уровень сложности:* задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте - 1 балл; ответ отличен от эталонного - 0 баллов.

*Повышенный уровень сложности:* задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, лишние символы в ответе отсутствуют - 2 балл; если на любой одной позиции ответа записан не тот символ, который представлен в эталоне ответа - 1 балл; во всех других случаях выставляется 0 баллов

**Комбинированные задания**

*Базовый уровень сложности:* задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, обоснование по смыслу соответствует эталонному (допускаются различные формулировки ответа, не искажающие его смысла) - 1 балл; ответ отличен от эталонного - 0 баллов.

*Повышенный уровень сложности:* задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, обоснование по смыслу соответствует эталонному (допускаются различные формулировки ответа, не искажающие его смысла) - 2 балла; дан верный ответ, обоснование отсутствует или приведено неверно – 1 балл; во всех остальных случаях - 0 баллов.

**Задания открытого типа**

*Повышенный уровень сложности:* ответ соответствует эталонному (допускаются различные формулировки ответа, не искажающие его смысла); правильно названы все запрашиваемые составляющие вопросы, даны верные обоснования - 2 балла; ответ имеет незначительные отклонения от эталонного, правильно названы на все запрашиваемые составляющие вопросы, но для названных даны верные обоснования - 1 балл; ответ значительно отличается от эталонного, имеются фактические ошибки, искажающие его смысл или ответ сформулирован неверно или не сформулирован - 0 баллов.

*Высокий уровень сложности:* магистрант демонстрирует умение применять знания в нестандартной ситуации, решать нетиповые задачи, приводит корректные обоснования и доказательства, ответ полный, в ответе отсутствуют фактические ошибки, изложение связное, структура прозрачная, логика изложения прослеживается - 3 балла; ответ значительно отличается от эталонного, имеются фактические ошибки, искажающие его смысл или ответ сформулирован неверно или не сформулирован - 0 баллов.

Итоговый балл за тест рассчитывается по формуле:

$$F = \frac{100}{K} * \left( \frac{x_1}{k_1} + \frac{x_2}{k_2} + \dots + \frac{x_n}{k_n} \right),$$

где  $F$  – итоговое количество баллов за тест,

$K$  – количество осваиваемых в рамках дисциплины компетенций,

$k_n$  – максимально возможное количество баллов за вопросы по компетенции,

$x_n$  – количество баллов, набранное магистрантом, за правильные ответы на вопросы по соответствующей компетенции.

Критерии оценивания тестирования приведены в Таблице 7.

Таблица 7

**Критерии оценивания письменной зачётной работы**

<b>Вид промежуточной аттестации</b>	<b>Показатели</b>	<b>Количество баллов</b>
Тест	81–100% правильных ответов	49–60
	61–80% правильных ответов	37–48
	41–60% правильных ответов	25–36
	0–40% правильных ответов	0–24

В результате промежуточного контроля знаний студенты получают аттестацию по дисциплине. На основании оценки обучающегося по итогам освоения дисциплины, выраженной в 100-балльной шкале, выставляется **зачет с оценкой** в соответствии с Таблицей 8.

Таблица 8  
**Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации**

<b>Форма промежуточной аттестации</b>	<b>Коды компетенций</b>	<b>Индикаторы компетенций (в соотв. с Таблицей 1)</b>	<b>Коды ЗУВ (в соотв. с Таблицей 1)</b>	<b>Оценка по итогам освоения дисциплины (в 100-балльной шкале)</b>	<b>Результаты текущего контроля</b>
<b>зачет с оценкой</b>	ОПК-2-Э ОПК-5-Э	ИД.ОПК-2.1.	З (ОПК-2)	81–100	Зачтено, отлично
		ИД.ОПК-2.2	У (ОПК-2)	61–80	Зачтено, хорошо
		ИД.ОПК-2.3	В (ОПК-2)	41–60	Зачтено, удовлетворительно
		ИД.ОПК-5.1. ИД.ОПК-5.3. ИД.ОПК-5.4. ИД.ОПК-5.5.	З (ОПК-5) У (ОПК-5) В (ОПК-5)	0–40	Не зачтено, неудовлетворительно

Результаты промежуточного контроля по дисциплине, выраженные в оценках «зачтено, удовлетворительно», «зачтено, хорошо», «зачтено, отлично» показывают уровень сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Данные, знания, экономика, нарративы» по направлению подготовки 38.04.01 Экономика (уровень магистратуры).

Результаты промежуточного контроля по дисциплине, выраженные в оценке «не зачтено, неудовлетворительно», показывают не сформированность у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Данные, знания, экономика, нарративы» по направлению подготовки 38.04.01 Экономика (уровень магистратуры).

## 7.4 Типовые задания к промежуточной аттестации

### 7.4.1. Перечень вопросов для подготовки к зачету

1. Типы статистических данных и типы моделей.
2. Линейная регрессионная модель. Формулировка задачи и используемые предпосылки.
3. Метод наименьших квадратов.
4. Коэффициент детерминации и скорректированный коэффициент детерминации.
5. Теорема Гаусса-Маркова.
6. Проверка значимости уравнения в целом. Проверка гипотез о коэффициентах.
7. Доверительный интервал для коэффициента.
8. Интерпретация коэффициентов модели.
9. Прогнозирование.
10. Фиктивные переменные и их использование.
11. Тесты Чоу на стабильность параметров.

12. Ошибки спецификации (пропущенные переменные, включение лишних переменных, выбор формы модели) и их последствия. RESET тест Рамсея.
13. Ослабление предпосылок классической линейной регрессионной модели: гетероскедастичность (тестирование гетероскедастичности, оценивание модели в условиях гетероскедастичности).
14. Ослабление предпосылок классической линейной регрессионной модели: автокорреляция (тестирование автокорреляции, оценивание модели в условиях автокорреляции).
15. Стохастические регрессоры. Эндогенность и метод инструментальных переменных. Тест Хаусмана.

#### **7.4.2. Пример тестового задания к зачету**

Компетенция: ОПК-2-Э

Уровень сложности: Базовый

Инструкция: Выберите один правильный вариант из предложенных.

Формулировка задания:

Что означает свойство несмешённости оценок МНК?

Варианты ответа:

- A. Ошибки имеют одинаковую дисперсию
- B. Среднее значение оценок коэффициентов совпадает с истинными значениями
- C. Остатки модели равны нулю
- D. Дисперсия ошибок минимальна

Поле для ответа:

--	--	--

Компетенция: ОПК-2-Э

Уровень сложности: Базовый

Инструкция: Выберите один правильный вариант из предложенных.

Формулировка задания:

Сумма остатков в модели линейной регрессии со свободным членом, оцененной методом наименьших квадратов всегда равна

Варианты ответа:

- A. 1
- B. 0
- C. Дисперсии ошибок
- D. Свободному члену

Поле для ответа:

--	--	--

Компетенция: ОПК-2-Э

Уровень сложности: Базовый

Инструкция: Выберите один правильный вариант из предложенных.

Формулировка задания:

В каких случаях следует использовать линейную регрессию с фиктивными переменными?

Варианты ответа:

- A. Когда объясняющая переменная категориальна
- B. Когда зависимая переменная категориальна
- C. При наличии гетероскедастичности
- D. При наличии автокорреляции

Поле для ответа:

--	--	--

Компетенция: ОПК-2-Э

Уровень сложности: Базовый

Инструкция: Выберите один правильный вариант из предложенных.

Формулировка задания:

Что представляет собой  $t$ -статистика для коэффициента при объясняющей переменной в регрессионном анализе?

Варианты ответа:

- A. Коэффициент детерминации
- B. Коэффициент корреляции регрессора и зависимой переменной
- C. Отношение оценки коэффициента к его стандартной ошибке
- D. Коэффициент, показывающий уровень мультиколлинеарности

Поле для ответа:

--	--	--

Компетенция: ОПК-5-Э

Уровень сложности: Базовый

Инструкция: Выберите один правильный вариант из предложенных.

Формулировка задания:

Какой из перечисленных ниже языков программирования чаще всего используется для построения регрессионных моделей в прикладной эконометрике?

Варианты ответа:

- A. HTML
- B. SQL
- C. R
- D. LaTeX

Поле для ответа:

--	--	--

Компетенция: ОПК-5-Э

Уровень сложности: Базовый

Инструкция: Выберите один правильный вариант из предложенных.

Формулировка задания:

Что делает функция write.csv() в R?

Варианты ответа:

- A. Импортирует таблицу из CSV-файла
- B. Записывает данные в CSV-файл
- C. Создает регрессионную модель
- D. Сортирует данные

Поле для ответа:

--	--	--

Компетенция: ОПК-5-Э

Уровень сложности: Базовый

Инструкция: Выберите один правильный вариант из предложенных.

Формулировка задания:

Какой тип компьютерных программ обеспечивает наилучшую поддержку визуального и аналитического анализа регрессионных моделей?

Варианты ответа:

- A. Текстовые редакторы
- B. Электронные таблицы
- C. Статистические пакеты
- D. Почтовые клиенты

Поле для ответа:

--	--	--

Компетенция: ОПК-5-Э

Уровень сложности: Базовый

Инструкция: Выберите один правильный вариант из предложенных.

Формулировка задания:

Что означает переменная Residual standard error в выводе функции summary() в R, применительно к модели линейной регрессии?

Варианты ответа:

- A. Среднее значение ошибки прогноза
- B. Оценка дисперсии ошибок
- C. Среднеквадратичная ошибка остатков
- D. Стандартная ошибка коэффициента

Поле для ответа:

--	--	--

Компетенция: ОПК-2-Э

Уровень сложности: Повышенный

Инструкция: Дайте краткий численный ответ. Ответ округлите до двух знаков после запятой.

Формулировка задания:

В модели  $y_i = \beta_0 + \beta_1 x_i + u_i$  получены:  $\hat{\beta}_1 = 1.8, SE = 0.6$ , где  $SE$  – стандартная ошибка коэффициента. Вычислите  $t$ -статистику и определите значим ли коэффициент на уровне 5%. Используйте двусторонний тест и нормальное приближение.

Поле для ответа: \_\_\_\_\_

Компетенция: ОПК-2-Э

Уровень сложности: Повышенный

Инструкция: Дайте краткий численный ответ. Ответ округлите до двух знаков после запятой.

Формулировка задания:

По выборке объема  $n = 20$  была оценена линейная регрессионная модель с  $k = 2$  регрессорами (помимо свободного члена). Коэффициент детерминации оказался равным  $R^2 = 0.81$ . Найдите скорректированный коэффициент детерминации.

Поле для ответа: \_\_\_\_\_

Компетенция: ОПК-2-Э

Уровень сложности: Повышенный

Инструкция: Дайте краткий численный ответ. Ответ обоснуйте.

Формулировка задания:

Найдите оценку коэффициента  $\beta_1$  в модели парной линейной регрессии  $y = \beta_0 + \beta_1 x + \varepsilon$ , если известно что  $\sum(x_i - \bar{x})^2 = 100$ ,  $\sum(x_i - \bar{x})(y_i - \bar{y}) = 250$ , где  $\bar{x}$  и  $\bar{y}$  – выборочные средние соответствующих переменных.

Поле для ответа: \_\_\_\_\_

Компетенция: ОПК-5-Э

Уровень сложности: Повышенный

Инструкция: Дайте развернутый ответ в свободной форме. Ответ должен включать интерпретацию коэффициентов с учётом их шкалы измерения.

Формулировка задания:

Вы оценили следующую модель:

$$health\_ex_i = \beta_0 + \beta_1 \cdot \log(income) + \beta_2 \cdot age + u$$

где

$health\_ex$  – индивидуальные ежемесячные расходы на здоровье (в рублях), По результатам  $income$  – доход (в рублях),  $age$  – возраст (в годах).

оценки получено:

$$\hat{\beta}_1 = 0.8$$

$$\hat{\beta}_2 = 0.04$$

Как правильно интерпретировать эти коэффициенты?

Поле для ответа: \_\_\_\_\_

Компетенция: ОПК-5-Э

Уровень сложности: Повышенный

Инструкция: Дайте краткий численный ответ. Ответ обоснуйте.

Формулировка задания:

В результате оценки регрессионной модели, включающей регрессор  $x_1$ , в статистическом пакете вы получили следующий результат:

Переменная Оценка коэффициента Стандартная ошибка

$x_1$	2.4	0.8
-------	-----	-----

Постройте 95% доверительный интервал для коэффициента при переменной  $x_1$ , если известно, что квантиль уровня 0.975 распределения Стьюдента с соответствующими степенями свободы равен 1.96.

Поле для ответа: \_\_\_\_\_

Компетенция: ОПК-5-Э

Уровень сложности: Повышенный

Инструкция: Выберите один правильный вариант из предложенных. Обоснуйте ответ.

Формулировка задания:

В статье приводится результат оценки линейной регрессии  $\hat{y} = 1.1 - 0.7 x$ . Как вы думаете,

что именно указано в скобках? Ответ обоснуйте. Варианты ответа:

- A.  $p$ -значения
- B.  $t$ -статистики
- C. Стандартные ошибки коэффициентов
- D. Показатели VIF для каждого коэффициента

Поле для ответа:

--	--	--

Обоснование \_\_\_\_\_

Компетенция: ОПК-2-Э

Уровень сложности: Высокий

Инструкция: Приведите развернутое рассуждение. Приведите пример и объясните интерпретацию коэффициентов.

Формулировка задания:

Что означает наличие произведения непрерывной и фиктивной переменных в регрессии? Приведите пример и объясните интерпретацию коэффициентов.

Поле для ответа: \_\_\_\_\_

Компетенция: ОПК-2-Э

Уровень сложности: Высокий

Инструкция: Приведите развернутое рассуждение.

Формулировка задания:

Ваш коллега утверждает, что модель с  $R^2 = 0.9$  всегда лучше модели с  $R^2 = 0.7$  пусть даже и с меньшим набором регрессоров. Когда модель с более низким  $R^2$  может быть предпочтительнее?

Поле для ответа: \_\_\_\_\_

Компетенция: ОПК-2-Э

Уровень сложности: Высокий

Инструкция: Прокомментируйте корректность утверждения и приведите объяснение, опираясь на выводы из оценки модели.

Формулировка задания:

Модель линейной регрессии имеет коэффициент детерминации  $R^2 = 0.95$ , но один из коэффициентов (при переменной  $X_1$ ) имеет  $p$ -значение = 0.55. Можно ли утверждать, что переменная  $X_1$  важна для модели?

Поле для ответа: \_\_\_\_\_

Компетенция: ОПК-5-Э

Уровень сложности: Высокий

Инструкция: Проанализируйте таблицу и выберите модель, которую следует предпочесть, обоснуйте выбор, ссылаясь на показатели. Ответ должен быть развёрнутым.

Формулировка задания:

В статистическом пакете вы оценили две модели зависимости потребительских расходов от дохода и других факторов. Программа вывела следующую информацию:

Модель	$R^2$	AIC	BIC	Число параметров
A	0.76	220.4	230.1	6
B	0.76	215.2	223.5	4

Какую модель следует предпочесть? Ответ поясните.

Поле для ответа: \_\_\_\_\_

Компетенция: ОПК-5-Э

Уровень сложности: Высокий

Инструкция: Проанализируйте результаты теста Хаусмана. Сделайте обоснованный вывод о предпочтительности модели. Ответ должен быть развёрнутым.

Формулировка задания:

Вы исследуете как количество сотрудников в фирме (employees) влияет на годовую выручку (revenue). Предполагается, что employees может быть эндогенной переменной (например, фирмы с высокой выручкой нанимают больше людей).

В качестве инструмента используется переменная regional\_wage (средняя региональная зарплата). Также в модель включена переменная sector (отрасль).

После оценки моделей МНК и 2SLS (двухшаговый МНК) вы провели тест Хаусмана. Статистический пакет привел следующие результаты:

- $\chi^2 = 6.32$  (значение статистики в teste Хаусмана),
- $p$ -значение = 0.012.

Какой вывод позволяет сделать тест Хаусмана? Какая из оценок, оценка МНК или оценка 2SLS, предпочтительна? Используйте уровень значимости 5%.

Поле для ответа: \_\_\_\_\_

Компетенция: ОПК-5-Э

Уровень сложности: Высокий

Инструкция: Выполните расчёт и дайте развёрнутый ответ. Ответ должен включать формулу, подстановку значений, результат вычислений и интерпретацию статистического критерия.

Формулировка задания:

В результате оценки двух моделей получены следующие суммы квадратов остатков:

- $RSS_1 = 150$  (базовая модель),
- $RSS_2 = 120$  (расширенная модель).
- Выборка содержит  $n = 50$  наблюдений.
- Базовая модель содержит  $k_1 = 2$  коэффициента, расширенная модель —  $k_2 = 4$  коэффициента.

Проверьте по F-критерию, является ли расширенная модель статистически лучше на уровне значимости 5%. Используйте критическое значение  $F_{0.05}(2, 46) = 3.20$ .

Поле для ответа: \_\_\_\_\_

## 7.5 Средства оценки индикаторов достижения компетенций

Таблица 9

### Средства оценки индикаторов достижения компетенций

Коды компетенций	Индикаторы компетенций (в соотв. с Таблицей 1)	Средства оценки (в соотв. с Таблицами 5, 7)
ОПК-2-Э	ИД.ОПК-2.1. ИД.ОПК-2.2 ИД.ОПК-2.3	домашнее задание (ДЗ), тест
ОПК-5-Э	ИД.ОПК-5.1. ИД.ОПК-5.3. ИД.ОПК-5.4. ИД.ОПК-5.5.	домашнее задание (ДЗ), тест

Таблица 10

### Описание средств оценки индикаторов достижения компетенций

Средства оценки (в соотв. с Таблицами 5, 7)	Рекомендованный план выполнения работы
Домашнее задание	Магистрант должен быть готовым в ходе подготовки и представления домашнего задания по темам дисциплины, выполнять следующие действия: — Для решения поставленных научно-исследовательских задач различного характера применяет продвинутые инструментальные методы экономического анализа

<b>Средства оценки (в соот. с Таблицами 5, 7)</b>	<b>Рекомендованный план выполнения работы</b>
	— Использует современные информационные технологии и программные средства при решении профессиональных задач.
Тест	Магистрант должен быть готовым в ходе подготовки к тесту, выполнять следующие действия: — Для решения поставленных научно-исследовательских задач различного характера применяет продвинутые инструментальные методы экономического анализа — Использует современные информационные технологии и программные средства при решении профессиональных задач.

## **8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА**

### **8.1. Основная литература**

1. Я. Р. Магнус, П. К. Катышев, А. А. Пересецкий. Эконометрика. Начальный курс: учебник / Академия народного хозяйства при Правительстве РФ. - 8-е изд. - Москва: Дело, 2007. - 503 с.
2. Грин У.Г. Эконометрический анализ. Книга 1 [Текст] / У. Г. Грин. - М.: ИД «Дело», 2016. - 760 с.
3. С. А. Айвазян, С. С. Иванова. Эконометрика: учеб. пособ.: рекомендовано УМО по образованию в области статистики в качестве учебного пособия для студентов высших учебных заведений, обучающихся по специальности 061700 "Статистика" и другим экономическим специальностям / - М. : Маркет DC, 2010. - 98 с.

### **8.2. Дополнительная литература**

1. Кэмерон, Колин. Микроэконометрика: методы и их применение [Текст]: учебник для вузов : в 2 кн. : перевод с английского / Э. К. Кэмерон, П. К. Триведи ; ред. пер. Б. Демешев ; Российская академия народного хозяйства и государственной службы при Президенте РФ. - М. : Дело, 2015. - (Академический учебник).

## **9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА**

### **9.1 Программное обеспечение**

1. OS Microsoft Windows (OVS OS Platform)
2. MS Office (OVS Office Platform)
3. Adobe Acrobat Professional 11.0 MLP AOO License RU
4. Adobe CS5.5 Design Standart Win IE EDU CLP
5. ABBYY FineReader 11 Corporate Edition
6. ABBYY Lingvo x5
7. Adobe Acrobat Reader DC /Pro – бесплатно
8. Google Chrome – бесплатно
9. Opera – бесплатно
10. Mozilla – бесплатно
11. VLC – бесплатно
12. Яндекс Браузер
13. R – бесплатно
14. Python – бесплатно

### **9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины:**

**Информационно-справочные системы**

- Гарант.Ру. Информационно-правовой портал: <http://www.garant.ru>
- Информационная система «Единое окно доступа к образовательным ресурсам»: <http://window.edu.ru/>
- Открытое образование. Ассоциация «Национальная платформа открытого образования»: <http://npoed.ru>
- Официальная Россия. Сервер органов государственной власти Российской Федерации: <http://www.gov.ru>
- Официальный интернет-портал правовой информации. Государственная система правовой информации: <http://pravo.gov.ru>
- Правовой сайт КонсультантПлюс: <http://www.consultant.ru/sys>
- Российское образование. Федеральный портал: <http://www.edu.ru>

**Профессиональные базы данных информационно-телекоммуникационной сети «Интернет»:**

- Национальная электронная библиотека НЭБ: <http://www.rusneb.ru>
- Президентская библиотека: <http://www.prlib.ru>
- Российская государственная библиотека: <http://www.rsl.ru/>
- Российская национальная библиотека: <http://www.nlr.ru/poisk/>

### **9.3 Лицензионные электронные ресурсы библиотеки Университета**

**Профессиональные базы данных:**

Полный перечень доступных обучающимся профессиональных баз данных представлен на официальном сайте Университета <https://eusp.org/library/electronic-resources, включая следующие базы данных>:

- eLIBRARY.RU** — Российский информационно-аналитический портал в области науки, технологии, медицины и образования, содержащий рефераты и полные тексты научных статей и публикаций, научометрическая база данных: <http://elibrary.ru>;
- Электронные журналы по подписке (текущие номера научных зарубежных журналов).

**Электронные библиотечные системы:**

- Znanium.com** – Электронная библиотечная система (ЭБС) – <http://znanium.com/>;
- Университетская библиотека онлайн – Электронная библиотечная система (ЭБС) – <http://biblioclub.ru/>

### **9.4 Электронная информационно-образовательная среда Университета**

Образовательный процесс по дисциплине поддерживается средствами электронной информационно-образовательной среды Университета, которая включает в себя электронный учебно-методический ресурс АНООВО «ЕУСПб» — образовательный портал LMS Sakai — [Sakai@EU](mailto:Sakai@EU), лицензионные электронные ресурсы библиотеки Университета, официальный сайт Университета (Европейский университет в Санкт-Петербурге [<https://eusp.org/>]), локальную сеть и корпоративную электронную почту Университета, и обеспечивает:

- доступ к учебным планам, рабочим программам дисциплин (модулей), практик и к изданиям электронных библиотечных систем и электронным образовательным ресурсам, указанным в рабочих программах;
- фиксацию хода образовательного процесса, результатов промежуточной аттестации и результатов освоения основной образовательной программы;
- формирование электронного портфолио обучающегося, в том числе сохранение работ обучающегося, рецензий и оценок за эти работы со стороны любых участников образовательного процесса;

- взаимодействие между участниками образовательного процесса, в том числе синхронное и (или) асинхронное взаимодействие посредством сети «Интернет» (электронной почты и т.д.).

Каждый обучающийся в течение всего периода обучения обеспечен индивидуальным неограниченным доступом к электронным ресурсам библиотеки Университета, содержащей издания учебной, учебно-методической и иной литературы по изучаемой дисциплине

## **10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА**

В ходе реализации образовательного процесса используются специализированные многофункциональные аудитории для проведения занятий лекционного типа, занятий семинарского типа (практических занятий, лабораторных работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованные специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Проведение занятий лекционного типа обеспечивается демонстрационным оборудованием.

Помещения для самостоятельной работы оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду организации.

Для лиц с ограниченными возможностями здоровья и инвалидов предоставляется возможность присутствия в аудитории вместе с ними ассистента (помощника). Для слабовидящих предоставляется возможность увеличения текста на экране ПК. Для самостоятельной работы лиц с ограниченными возможностями здоровья в помещении для самостоятельной работы организовано одно место (ПК) с возможностями бесконтактного ввода информации и управления компьютером (специализированное лицензионное программное обеспечение – Camera Mouse, веб камера). Библиотека университета предоставляет удаленный доступ к электронным ресурсам библиотеки Университета с возможностями для слабовидящих увеличения текста на экране ПК. Лица с ограниченными возможностями здоровья могут при необходимости воспользоваться имеющимся в университете креслом-коляской. В учебном корпусе имеется адаптированный лифт. На первом этаже оборудован специализированный туалет. У входа в здание университета для инвалидов оборудована специальная кнопка, входная среда обеспечена информационной доской о режиме работы университета, выполненной рельефно-точечным тактильным шрифтом (азбука Брайля).

**ПРИЛОЖЕНИЕ 1**

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ**  
**«Основы регрессионного анализа»**

## **ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ**

### **1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации**

Информация о содержании и процедуре текущего контроля успеваемости, методике оценивания знаний, умений и навыков обучающегося в ходе текущего контроля доводятся научно-педагогическими работниками Университета до сведения обучающегося на первом занятии по данной дисциплине.

Текущий контроль предусматривает подготовку магистрантов к каждому практическому занятию, активное слушание на лекциях, выполнение магистрантами домашних заданий. Магистрант должен присутствовать на занятиях, отвечать на поставленные вопросы, показывая, что прочитал разбираемую литературу, представлять содержательные реплики по темам обсуждения.

Текущий контроль проводится в форме выполнения магистрантами домашних заданий, демонстрирующих степень знакомства магистрантов с дополнительной литературой.

Таблица 1

**Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации**

Наименование разделов	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Тема 1. Множественная регрессия – основные понятия	1. ОПК-2-Э ОПК-5-Э	ИД.ОПК-2.1. ИД.ОПК-2.2 ИД.ОПК-2.3 ИД.ОПК-5.1. ИД.ОПК-5.3. ИД.ОПК-5.4. ИД.ОПК-5.5.	З (ОПК-2) У (ОПК-2) В (ОПК-2) З (ОПК-5) У (ОПК-5) В (ОПК-5)	Домашнее задание	зачтено/ не зачтено
Тема 2. Ослабление предпосылок классической линейной регрессионной модели.	ОПК-2-Э ОПК-5-Э	ИД.ОПК-2.1. ИД.ОПК-2.2 ИД.ОПК-2.3 ИД.ОПК-5.1. ИД.ОПК-5.3. ИД.ОПК-5.4. ИД.ОПК-5.5.	З (ОПК-2) У (ОПК-2) В (ОПК-2) З (ОПК-5) У (ОПК-5) В (ОПК-5)	Домашнее задание	зачтено/ не зачтено

При освоении дисциплины каждая из форм текущего контроля оценивается с использованием 100-балльной шкалы с последующим переводом в бинарную систему для получения результатов текущего контроля, фиксирующих ход образовательного процесса, согласно Таблице 2. Максимально возможное количество баллов для домашней работы – 40 баллов.

Таблица 2

**Критерии оценивания**

Формы текущего контроля успеваемости	Описание	Показатели	Количество баллов по 100-балльной шкале	Результаты текущего контроля
Домашнее задание	Домашнее задание представляет собой набор задач по пройденным темам к моменту отчетного занятия. Задачи,	Магистрант глубоко и прочно усвоил программный материал; исчерпывающе, последовательно, четко и логически стройно излагает	91-100% от максимального возможного	зачтено

<b>Формы текущего контроля успеваемости</b>	<b>Описание</b>	<b>Показатели</b>	<b>Количество баллов по 100-балльной шкале</b>	<b>Результаты текущего контроля</b>
	предлагаемые студентам, дают им представление о сложности заданий, предлагаемых на контрольной работе. Студенты решают задачи дома и приносят их (или отправляют преподавателю по электронной почте) к поставленному сроку (дедлайну).	его в письменной домашней работе, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами знаний.	о числа баллов	
		Магистрант твердо знает материал, грамотно и по существу излагает его в письменной домашней работе, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения.	51-90% от максимального возможного числа баллов	
		Магистрант имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в ответах письменной домашней работы.	25-50% от максимального возможного числа баллов	
		Магистрант не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями отвечает на вопросы письменной домашней работы.	1-24% от максимального возможного числа баллов	не засчитено

## 2 Контрольные задания для текущей аттестации

### Материал домашних заданий

#### Тема 1: Множественная регрессия – основные понятия

1. Найдите оценку метода наименьших квадратов параметра  $b$  в регрессии  $y = b + \varepsilon$ , где  $y$  – зависимая переменная,  $b$  – константа подлежащая оценке,  $\varepsilon$  – ошибки. Чему равен коэффициент детерминации в оцененной модели?

2. Изменение единиц измерения зависимой переменной в регрессионной модели  $y = b_0 + b_1x_1 + b_2x_2 + \varepsilon$  НЕ приведёт к изменению (выберите верный вариант ответа)

- оценки коэффициента  $b_0$
- оценки коэффициентов  $b_1$  и  $b_2$

- коэффициента детерминации
- общей суммы квадратов остатков

3. Пусть QCOFFEE – спрос на кофе, PCOFFEE – цена кофе, PTEA – цена чая, PCREAMER – цена сливок, INC – доход потребителей.

**Модель 1:** МНК, использованы наблюдения 1-21

Зависимая переменная: lnQCOFFEE

	Коэффициент	Ст. ошибка	t- статистика	P-значение
const	-2,43973	1,59313	-1,5314	0,14520
lnPCOFFEE	-3,31253	0,688323	-4,8125	???????
lnPTEA	1,5942	0,571989	???????	0,01319
lnPCREAMER	0,0693744	????????	0,1117	0,91247
lnINC	????????	0,0993055	3,0110	0,00829
Среднее зав. перемен	1,491580	Ст. откл. зав. перемен		0,504265
Сумма кв. остатков	1,320197	Ст. ошибка модели		0,287250
R-квадрат	0,740408	Испр. R-квадрат		??????????
F(4, 16)	?????????	P-значение (F)		0,000143
Лог.правдоподобие	-0,746924	Крит. Акаике		11,49385
Крит. Шварца	16,71646	Крит. Хеннана-Куинна		12,62729

- 1) Восстановите цифры, заменённые на знак вопроса.
- 2) Выпишите оцененное уравнение.
- 3) Проверьте значимость уравнения в целом.
- 4) Проверьте значимость отдельных коэффициентов.
- 5) Интерпретируйте коэффициенты.
- 6) Является ли спрос на кофе эластичным?
- 7) Является ли чай субститутом кофе?
- 8) Является ли кофе нормальным товаром?
- 9) Является ли кофе предметом роскоши?
- 10) Являются ли сливки комплементарным товаром?

4. В файле tropic.csv содержатся данные для 84 стран мира:

GROWTH – среднегодовой темп экономического роста за 1965–1990 годы,  
LOGGDP65 – логарифм ВНП на душу населения в 1965 году,  
LND100km – доля территории, находящейся не далее 100 км от побережья,  
POP100km – доля населения, проживающего не далее 100 км от побережья,  
TROPICAR – доля территории, находящейся в тропиках,  
TROPOPOP – доля населения, проживающего в тропиках,  
OPEN – доля лет в 1965–1990 гг., в течение которых страна была открыта для международной торговли,  
POLINST – индекс качества политических институтов (от 0 – низкое качество, до 10 – высокое качество).

- a) Оцените зависимость темпа экономического роста от всех остальных переменных с помощью множественной линейной регрессии.
- b) Как можно интерпретировать коэффициенты при объясняющих переменных в этой регрессии?

- c) Какие из коэффициентов статистически незначимы?
- d) Являются ли незначимые коэффициенты незначимыми **совокупно** (в смысле F-теста)?
- e) Можно ли по результатам оценки утверждать, что более богатые страны (по уровню ВНП на душу населения в 1965 году) росли медленнее?
- f) Можно ли утверждать, что экономический рост был ниже в странах с большей долей территории или населения в тропиках?

5. Файл `real_estate_price_size_year_view.csv` содержит данные:

`price` – цена недвижимости  
`size` – размер недвижимости (в квадратных метрах)  
`year` – год постройки  
`view` – наличие вида на море (текстовая переменная «`Sea view`» или «`No sea view`»)

- a) Загрузите данные и создайте дамми-переменную для переменной `view`, где 1 соответствует «`Sea view`», а 0 — «`No sea view`».
- b) Постройте модель множественной линейной регрессии, где зависимая переменная — `price`, а независимые переменные — `size`, `year` и дамми-переменная `view`.
- c) Интерпретируйте коэффициенты модели. Особое внимание уделите коэффициенту при дамми-переменной `view`.
- d) Проверьте значимость коэффициентов модели с использованием t-тестов.
- e) Постройте график зависимости цены от размера недвижимости, разделив данные на группы по наличию вида на море
- f) Проверьте, зависит ли эффект размера недвижимости на её цену от наличия вида на море, включив в модель взаимодействие между переменными `size` и `view`.

## **Тема 2: Ослабление предпосылок классической линейной регрессионной модели.**

- 1. Что из следующего является следствием гетероскедастичности?
  - a) МНК-оценки коэффициентов регрессии смещены.
  - b) МНК-оценки коэффициентов регрессии несостоительны.
  - c) МНК-оценки коэффициентов регрессии неэффективны.
  - d) МНК-оценка матрицы ковариаций коэффициентов регрессии является смещённой и несостоительной.
  - e) Применение стандартных тестов (Стьюдента, Фишера) некорректно.
- 2. Что из следующего является следствием автокорреляции?
  - a) МНК-оценки коэффициентов регрессии смещены.
  - b) МНК-оценки коэффициентов регрессии несостоительны.
  - c) МНК-оценки коэффициентов регрессии неэффективны.
  - d) МНК-оценка матрицы ковариаций коэффициентов регрессии является смещённой и несостоительной.
  - e) Применение стандартных тестов (Стьюдента, Фишера) некорректно.

2. Файл `financial_wealth13.csv` содержит данные
- `nettfa` – общие чистые финансовые активы человека, в тысячах денежных единиц  
`inc` – доход человека, в тысячах денежных единиц

age – возраст, лет

agesq – возраст в квадрате

e401k = 1, если человек участвует в накопительном пенсионном плане, и = 0, если не участвует.

Оцените уравнение

$$\text{NETTFA} = C(1) * \text{INC} + C(2) * \text{age} + C(3) * \text{agesq} + C(4) * \text{e401k} + \varepsilon$$

- a) Протестируйте его на наличие гетероскедастичности всеми известными вам способами.
- b) Если гетероскедастичность будет обнаружена, предложите способы того, как решить эту проблему и реализуйте это.
- c) Прокомментируйте свои ответы и решения.

4. Используя данные Росстата и Банка России за период с 2008 по 2024 год, проанализируйте зависимость потребления домашних хозяйств от располагаемых доходов и процентной ставки.

Переменные:

cons — расходы на конечное потребление домашних хозяйств, млрд руб.

income — располагаемые денежные доходы населения, млрд руб.

rate — средняя процентная ставка по кредитам физическим лицам

- a) Постройте линейную регрессионную модель:

$$\text{cons} = \beta_0 + \beta_1 * \text{income} + \beta_2 * \text{rate} + \varepsilon$$

- b) Проверьте наличие автокорреляции первого порядка в остатках модели с помощью теста Дарбина–Уотсона.
- c) Обсудите, как автокорреляция влияет на:
  - оценку коэффициента;
  - корректность стандартных ошибок;
  - доверительные интервалы и t-статистики.

4. Файл *fertility.csv* содержит следующие данные:

HOURSWM – количество рабочих часов женщины в неделю (зависимая переменная),

MOREKIDS – дамми: 1, если у женщины более двух детей; 0 – иначе

AGEFSTM – возраст матери на момент рождения первого ребёнка

MULTI2ND – дамми: 1, если второй и третий ребёнок — близнецы

BOY1ST Дамми: 1, если первый ребёнок — мальчик

BOY2ND Дамми: 1, если второй ребёнок — мальчик

- a. Оцените МНК-регрессию hourswm на morekids и agefstm.
- b. Интерпретируйте коэффициент при переменной morekids.
- c. Есть ли основания считать, что этот коэффициент смещен и нам нужны инструментальные переменные для morekids?

- d. Постройте переменные two\_girls — оба первых ребёнка девочки, two\_boys — оба первых ребёнка мальчики.
- e. Обсудите, могут ли multi2nd, two\_girls и two\_boys выступать хорошими инструментами для morekids. Учитывайте два ключевых свойства инструментов:
  - релевантность;
  - экзогенность.
- f. Проведите тест на слабость инструментов.
- g. Проведите тест Хаусмана вручную.
- h. Если хотя бы один из инструментов выглядит слабым или подозрительным по смыслу — уберите его и повторите оценку.
- i. Если все инструменты проходят ваши проверки — сохраните исходную спецификацию.
- j. Сравните оценки параметров по МНК и IV-моделям. Как изменилась оценка при morekids? Как вы это интерпретируете?

### **3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации**

**Форма промежуточной аттестации – зачет с оценкой в форме тестирования.**

Перед зачетом с оценкой проводятся консультации, на которых преподаватель отвечает на вопросы магистрантов.

В результате промежуточного контроля знаний студенты получают аттестацию по дисциплине.

При выставлении оценки по итогам освоения дисциплины учитываются как результаты тестирования, так и текущего контроля успеваемости. Максимальная оценка по итогам освоения дисциплины составляет 100 баллов и определяется суммой баллов, полученных обучающимся за домашние задания (40 баллов максимально) и за тестирование (60 баллов максимально).

При вычислении оценки обучающегося по итогам освоения дисциплины используется процедура округления до целого.

В рамках тестирования студентам будет необходимо решить 20 заданий базовой, повышенной и высокой сложности, из которых 8 заданий являются заданиями базовой сложности, 6 заданий повышенной сложности и 6 заданий высокой сложности. Задания распределяются между компетенциями в равной пропорции для всех уровней сложности. В базовых заданиях проверяются знания студента по прошедшему курсу, в остальных – способность применять знания и логически обосновать свой выбор.

Тест оценивается в баллах в соответствии со следующими критериями:

#### **Задания закрытого типа**

**Базовый уровень сложности:** задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте - 1 балл; ответ отличен от эталонного - 0 баллов.

**Повышенный уровень сложности:** задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, лишние символы в ответе отсутствуют - 2 балла; если на любой одной позиции ответа записан не тот символ, который представлен в эталоне ответа - 1 балл; во всех других случаях выставляется 0 баллов

#### **Комбинированные задания**

*Базовый уровень сложности:* задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, обоснование по смыслу соответствует эталонному (допускаются различные формулировки ответа, не искажающие его смысла) - 1 балл; ответ отличен от эталонного - 0 баллов.

*Повышенный уровень сложности:* задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, обоснование по смыслу соответствует эталонному (допускаются различные формулировки ответа, не искажающие его смысла) - 2 балла; дан верный ответ, обоснование отсутствует или приведено неверно – 1 балл; во всех остальных случаях - 0 баллов.

### **Задания открытого типа**

*Повышенный уровень сложности:* ответ соответствует эталонному (допускаются различные формулировки ответа, не искажающие его смысла); правильно названы все запрашиваемые составляющие вопросы, даны верные обоснования - 2 балла; ответ имеет незначительные отклонения от эталонного, правильно названы на все запрашиваемые составляющие вопросы, но для названных даны верные обоснования - 1 балл; ответ значительно отличается от эталонного, имеются фактические ошибки, искажающие его смысл или ответ сформулирован неверно или не сформулирован - 0 баллов.

*Высокий уровень сложности:* магистрант демонстрирует умение применять знания в нестандартной ситуации, решать нетиповые задачи, приводит корректные обоснования и доказательства, ответ полный, в ответе отсутствуют фактические ошибки, изложение связное, структура прозрачная, логика изложения прослеживается - 3 балла; ответ значительно отличается от эталонного, имеются фактические ошибки, искажающие его смысл или ответ сформулирован неверно или не сформулирован - 0 баллов.

Итоговый балл за тест рассчитывается по формуле:

$$F = \frac{100}{K} * \left( \frac{x_1}{k_1} + \frac{x_2}{k_2} + \dots + \frac{x_n}{k_n} \right),$$

где  $F$  – итоговое количество баллов за тест,  
 $K$  – количество осваиваемых в рамках дисциплины компетенций,  
 $k_n$  – максимально возможное количество баллов за вопросы по компетенции,  
 $x_n$  – количество баллов, набранное магистрантом, за правильные ответы на вопросы по соответствующей компетенции.

Критерии оценивания тестирования приведены в Таблице 3.

Таблица 3

### **Критерии оценивания письменной зачётной работы**

Вид промежуточной аттестации	Показатели	Количество баллов
Тест	81–100% правильных ответов	49-60
	61–80% правильных ответов	37-48
	41–60% правильных ответов	25–36
	0–40% правильных ответов	0–24

В результате промежуточного контроля знаний студенты получают аттестацию по дисциплине. На основании оценки обучающегося по итогам освоения дисциплины, выраженной в 100-балльной шкале, выставляется **зачет с оценкой** в соответствии с Таблицей 4.

Таблица 4

### **Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации**

Форма промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соотв. с Таблицей 1)	Коды ЗУВ (в соотв. с Таблицей 1)	Оценка по итогам освоения дисциплины (в 100-балльной шкале)	Результаты текущего контроля
зачет оценкой	с ОПК-2-Э ОПК-5-Э	ИД.ОПК-2.1.	З (ОПК-2)	81–100	Зачтено, отлично
		ИД.ОПК-2.2	У (ОПК-2)	61–80	Зачтено, хорошо
		ИД.ОПК-2.3	В (ОПК-2)	41–60	Зачтено, удовлетворительно
		ИД.ОПК-5.1. ИД.ОПК-5.3. ИД.ОПК-5.4. ИД.ОПК-5.5.	З (ОПК-5) У (ОПК-5) В (ОПК-5)	0–40	Не зачтено, неудовлетворительно

Результаты промежуточного контроля по дисциплине, выраженные в оценках «зачтено, удовлетворительно», «зачтено, хорошо», «зачтено, отлично» показывают уровень сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Данные, знания, экономика, нарративы» по направлению подготовки 38.04.01 Экономика (уровень магистратуры).

Результаты промежуточного контроля по дисциплине, выраженные в оценке «не зачтено, неудовлетворительно», показывают не сформированность у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Данные, знания, экономика, нарративы» по направлению подготовки 38.04.01 Экономика (уровень магистратуры).

## 4 Задания к промежуточной аттестации

### 4.1. Перечень вопросов для подготовки к зачету

1. Типы статистических данных и типы моделей.
2. Линейная регрессионная модель. Формулировка задачи и используемые предпосылки.
3. Метод наименьших квадратов.
4. Коэффициент детерминации и скорректированный коэффициент детерминации.
5. Теорема Гаусса-Маркова.
6. Проверка значимости уравнения в целом. Проверка гипотез о коэффициентах.
7. Доверительный интервал для коэффициента.
8. Интерпретация коэффициентов модели.
9. Прогнозирование.
10. Фиктивные переменные и их использование.
11. Тесты Чоу на стабильность параметров.
12. Ошибки спецификации (пропущенные переменные, включение лишних переменных, выбор формы модели) и их последствия. RESET тест Рамсея.
13. Ослабление предпосылок классической линейной регрессионной модели: гетероскедастичность (тестирование гетероседастичности, оценивание модели в условиях гетероскедастичности).
14. Ослабление предпосылок классической линейной регрессионной модели: автокорреляция (тестирование автокорреляции, оценивание модели в условиях автокорреляции).
15. Стохастические регрессоры. Эндогенность и метод инструментальных переменных. Тест Хаусмана.

## **4.2. Банк тестовых заданий**

**ОПК-2-Э: Способен применять продвинутые инструментальные методы экономического анализа в прикладных и (или) фундаментальных исследованиях**

**ОПК-5-Э: Способен использовать современные информационные технологии и программные средства при решении профессиональных задач**

**Вопросы на выбор одного правильного ответа**

**Компетенция ОПК-2-Э**

**Задания закрытого типа**

**Базовый уровень сложности**

### **Задание 1**

**Инструкция:** Прочитайте задание и выберите правильный ответ

Что означает наличии мультиколлинеарности в модели множественной регрессии?

- A. Гомоскедастичность
- B. Автокорреляцию
- C. Плохую объясняющую способность модели
- D. Сильную коррелированность регрессоров

**Поле для ответа:**

--	--	--

### **Задание 2**

**Инструкция:** Прочитайте задание и выберите правильный ответ

Что такое  $R^2$  в модели линейной регрессии?

- A. Коэффициент автокорреляции
- B. Показатель наличия гетероскедастичности
- C. Доля объяснённой дисперсии зависимой переменной
- D. Показатель наличия мультиколлинеарности

**Поле для ответа:**

--	--	--

### **Задание 3**

**Инструкция:** Прочитайте задание и выберите правильный ответ

Какой из нижеперечисленных методов позволяет корректно оценить стандартные ошибки в линейной регрессии при наличии гетероскедастичности?

- A. Ridge регрессия
- B. Метод главных компонент
- C. Робастная оценка стандартных ошибок
- D. Lasso регрессия

**Поле для ответа:**

--	--	--

#### Задание 4

**Инструкция:** Прочтите задание и выберите правильный ответ

Что означает значимость коэффициента при переменной в регрессионной модели?

- A. Что коэффициент точно положителен
- B. Что коэффициент статистически отличен от нуля
- C. Что коэффициент статистически не отличим от нуля
- D. Что ошибка модели минимальна

**Поле для ответа:**

--	--	--

#### Задание 5

**Инструкция:** Прочтите задание и выберите правильный ответ

Какой из нижеперечисленных признаков указывает на наличие автокорреляции остатков в регрессии?

- A. Высокое значение коэффициента детерминации
- B. Наличие значимых коэффициентов при регрессорах
- C. Близкие к нулю значения статистики в тесте Уайта (White)
- D. Близкие к нулю значения статистики в тесте Дарбина – Уотсона (Durbin-Watson)

**Поле для ответа:**

--	--	--

#### Задание 6

**Инструкция:** Прочтите задание и выберите правильный ответ

Какое из следующих действий позволяет устраниить полную мультиколлинеарность?

- A. Уменьшение объема выборки
- B. Исключение одного из регрессоров
- C. Использование робастных стандартных ошибок
- D. Добавление новых регрессоров

**Поле для ответа:**

--	--	--

#### Задание 7

**Инструкция:** Прочтите задание и выберите правильный ответ

Какова цель анализа остатков в регрессионной модели?

- A. Проверка отсутствия эндогенности
- B. Проверка значимости коэффициентов

- С. Проверка предпосылок модели линейной регрессии  
D. Проверка отсутствия мультиколлинеарности

**Поле для ответа:**

--	--	--

**Задание 8**

**Инструкция:** Прочтите задание и выберите правильный ответ

Как интерпретируется коэффициент при логарифмированной объясняющей переменной в лог-линейной модели (модели, в которой зависимая переменная и регрессоры взяты в логарифмах)?

- A. Эластичность  
B. Уровневое изменение  
C. Автокорреляция  
D. Ковариация

**Поле для ответа:**

--	--	--

**Задание 9**

**Инструкция:** Прочтите задание и выберите правильный ответ

Что представляет собой  $t$ -статистика для коэффициента при объясняющей переменной в регрессионном анализе?

- A. Коэффициент детерминации  
B. Коэффициент корреляции регрессора и зависимой переменной  
C. Отношение оценки коэффициента к его стандартной ошибке  
D. Коэффициент, показывающий уровень мультиколлинеарности

**Поле для ответа:**

--	--	--

**Задание 10**

**Инструкция:** Прочтите задание и выберите правильный ответ

В каких случаях следует использовать линейную регрессию с фиктивными переменными?

- A. Когда объясняющая переменная категориальна  
B. Когда зависимая переменная категориальна  
C. При наличии гетероскедастичности  
D. При наличии автокорреляции

**Поле для ответа:**

--	--	--

### **Задание 11**

**Инструкция:** Прочтите задание и выберите правильный ответ  
Какой метод можно использовать для диагностики нормальности остатков в модели линейной регрессии?

- A. Тест Уайта
- B. Тест Хархе – Бера
- C. Тест Дарбина-Уотсона
- D. Тест Стьюдента

**Поле для ответа:**

--	--	--

### **Задание 12**

**Инструкция:** Прочтите задание и выберите правильный ответ  
Сумма остатков в модели линейной регрессии со свободным членом, оцененной методом наименьших квадратов всегда равна

- A. 1
- B. 0
- C. Дисперсии ошибок
- D. Свободному члену

**Поле для ответа:**

--	--	--

### **Задание 13**

**Инструкция:** Прочтите задание и выберите правильный ответ  
Что означает свойство несмешённости оценок МНК?

- A. Ошибки имеют одинаковую дисперсию
- B. Среднее значение оценок коэффициентов совпадает с истинными значениями
- C. Остатки модели равны нулю
- D. Дисперсия ошибок минимальна

**Поле для ответа:**

--	--	--

### **Вопросы на выбор одного правильного ответа**

**Компетенция ОПК-5-Э**

**Задания закрытого типа**

**Базовый уровень сложности**

### **Задание 1**

**Инструкция:** Прочтите задание и выберите правильный ответ  
Какой из перечисленных ниже языков программирования чаще всего используется для построения регрессионных моделей в прикладной эконометрике?  
A. HTML  
B. SQL

- C. R
- D. LaTeX

**Поле для ответа:**

--	--	--

### Задание 2

**Инструкция:** Прочтайте задание и выберите правильный ответ  
Какой R-пакет используется для оценки моделей линейной регрессии?

- A. ggplot2
- B. lmtest
- C. dplyr
- D. lm

**Поле для ответа:**

--	--	--

### Задание 3

**Инструкция:** Прочтайте задание и выберите правильный ответ  
Какой тип графика помогает визуально выявить гетероскедастичность?  
A. Boxplot зависимой переменной  
B. Гистограммы регрессоров  
C. График остатков в модели  
D. График автокорреляционной функции

**Поле для ответа:**

--	--	--

### Задание 4

**Инструкция:** Прочтайте задание и выберите правильный ответ  
Что делает функция summary() в R при применении к модели линейной регрессии?  
A. Отображает график  
B. Строит модель  
C. Показывает все основные статистики, связанные с моделью линейной регрессии  
D. Загружает данные

**Поле для ответа:**

--	--	--

### Задание 5

**Инструкция:** Прочтайте задание и выберите правильный ответ  
Что означает строка кода в R: model <- lm(y ~ x1 + x2, data = df)?

- A. Сохраняет данные
- B. Создает таблицу корреляции
- C. Оценивает линейную регрессионную модель
- D. Очищает пропущенные значения

**Поле для ответа:**

--	--	--

**Задание 6**

**Инструкция:** Прочтите задание и выберите правильный ответ  
Какой из перечисленных ниже форматов файла чаще других используется для импорта данных в R?

- A. .exe
- B. .csv
- C. .jpg
- D. .zip

**Поле для ответа:**

--	--	--

**Задание 7**

**Инструкция:** Прочтите задание и выберите правильный ответ  
Какой тип данных требует в R преобразования в factor перед использованием в регрессии?  
A. Категориальная переменная  
B. Непрерывная переменная  
C. Временная переменная  
D. Логарифм зависимой переменной

**Поле для ответа:**

--	--	--

**Задание 8**

**Инструкция:** Прочтите задание и выберите правильный ответ  
Что означает переменная Residual standard error в выводе функции summary() в R, применительно к модели линейной регрессии?  
A. Среднее значение ошибки прогноза  
B. Оценка дисперсии ошибок  
C. Среднеквадратичная ошибка остатков  
D. Стандартная ошибка коэффициента

**Поле для ответа:**

--	--	--

### **Задание 9**

**Инструкция:** Прочтите задание и выберите правильный ответ

Что делает функция write.csv() в R?

- A. Импортирует таблицу из CSV-файла
- B. Записывает данные в CSV-файл
- C. Создает регрессионную модель
- D. Сортирует данные

**Поле для ответа:**

--	--	--

### **Задание 10**

**Инструкция:** Прочтите задание и выберите правильный ответ

Какая из перечисленных ниже задач наиболее типична при использовании ИТ-средств в регрессионном анализе?

- A. Установление причинно-следственных связей без данных
- B. Визуализация взаимосвязей между переменными
- C. Хранение бумажных архивов
- D. Ручной расчет среднеквадратичной ошибки

**Поле для ответа:**

--	--	--

### **Задание 11**

**Инструкция:** Прочтите задание и выберите правильный ответ

Какой тип компьютерных программ обеспечивает наилучшую поддержку визуального и аналитического анализа регрессионных моделей?

- A. Текстовые редакторы
- B. Электронные таблицы
- C. Статистические пакеты
- D. Почтовые клиенты

**Поле для ответа:**

--	--	--

### **Задание 12**

**Инструкция:** Прочтите задание и выберите правильный ответ

Что из нижеперечисленного не может обеспечить использование программного обеспечения в регрессионном анализе?

- A. Повышение точности расчетов
- B. Быстрая обработка больших объемов данных
- C. Устранение необходимости в интерпретации результатов
- D. Возможность автоматизации анализа

**Поле для ответа:**

--	--	--

**Задание 13**

**Инструкция:** Прочтите задание и выберите правильный ответ

Что из нижеперечисленного относится к этапу подготовки данных в регрессионном анализе с использованием программных средств?

- A. Подбор теоретической модели
- B. Формирование гипотез
- C. Очистка и кодирование данных
- D. Интерпретация коэффициентов модели

**Поле для ответа:**

--	--	--

**Вопросы на выбор нескольких правильных ответов**

**Компетенция ОПК-2-Э**

**Задания закрытого типа**

**Повышенный уровень сложности**

**Задание 1**

**Инструкция:** Прочтите задание и выберите все правильные ответы

Какие из приведенных ниже предпосылок используются в теореме Гаусса-Маркова?

- A. Линейность регрессионной функции
- B. Нормальность распределения объясняющих переменных
- C. Отсутствие автокорреляции ошибок
- D. Постоянная дисперсия ошибок (гомоскедастичность)

**Поле для ответа:**

--	--	--

**Задание 2**

**Инструкция:** Прочтите задание и выберите все правильные ответы

Что может указывать на наличие мультиколлинеарности в модели?

- A. Высокие значения VIF (variance inflation factor)
- B. Значения коэффициентов регрессии, близкие к нулю
- C. Значительное изменение оценок коэффициентов при добавлении/удалении одного из наблюдений
- D. Высокая коррелированность регрессоров

**Поле для ответа:**

--	--	--

### **Задание 3**

**Инструкция:** Прочтите задание и выберите все правильные ответы  
Какие из следующих утверждений соответствуют интерпретации коэффициента при фиктивной переменной в модели с константой?

- A. Разница в среднем уровне зависимой переменной между группами
- B. Процентное изменение зависимой переменной
- C. Коэффициент корреляции между переменными
- D. Доля наблюдений, относящихся к группе

**Поле для ответа:**

--	--	--

### **Задание 4**

**Инструкция:** Прочтите задание и выберите все правильные ответы  
Какие статистики обычно используются для оценки значимости регрессионной модели и ее отдельных коэффициентов?

- A. t-статистика
- B. F-статистика
- C. VIF (variance inflation factor)
- D. Коэффициент корреляции между регрессорами

**Поле для ответа:**

--	--	--

### **Задание 5**

**Инструкция:** Прочтите задание и выберите все правильные ответы  
Какие методы можно использовать для обнаружения гетероскедастичности?

- A. Тест Бреуша-Пагана
- B. Тест Стьюдента
- C. Тест Дарбина-Уотсона
- D. Тест Уайта

**Поле для ответа:**

--	--	--

### **Задание 6**

**Инструкция:** Прочтите задание и выберите все правильные ответы  
Как обычно влияет (частичная) мультиколлинеарность на оценки МНК?

- A. Делает оценки коэффициентов смещёнными
- B. Увеличивает стандартные ошибки оценок
- C. Уменьшает количество наблюдений
- D. Увеличивает R<sup>2</sup>

**Поле для ответа:**

--	--	--

### Задание 7

**Инструкция:** Прочтайте задание и выберите все правильные ответы

Что может являться следствием исключения существенной переменной из линейной регрессионной модели?

- A. Повышение дисперсии ошибок
- B. Увеличение коэффициента детерминации
- C. Снижение точности прогнозов
- D. Оценки коэффициентов модели несостоительны

**Поле для ответа:**

--	--	--

### Задание 8

**Инструкция:** Прочтайте задание и выберите все правильные ответы

При наличии автокорреляции ошибок каких свойств лишаются оценки МНК?

- A. Несмешённости
- B. Состоительности
- C. Эффективности (в классе линейных несмешенных оценок)
- D. Интерпретируемости

**Поле для ответа:**

--	--	--

### Задание 9

**Инструкция:** Прочтайте задание и выберите все правильные ответы

Какие последствия могут возникнуть при наличии гетероскедастичности в модели?

- A. Смешённость оценок коэффициентов
- B. Неверные стандартные ошибки
- C. Недостоверные t-статистики
- D. Смещение прогнозов

**Поле для ответа:**

--	--	--

### Задание 10

**Инструкция:** Прочтайте задание и выберите все правильные ответы

Какие признаки указывают на наличие ошибки спецификации?

- A. Высокое значение коэффициента детерминации
- B. Преимущественно отрицательные значения остатков
- C. Значимые коэффициенты
- D. Отврежение нулевой гипотезы в тесте RESET

**Поле для ответа:**

--	--	--

**Задание 11**

**Инструкция:** Прочтайте задание и выберите все правильные ответы

Какие признаки могут указывать на наличие мультиколлинеарности в модели?

- A. Высокие значения  $R^2$ , однако, коэффициенты в модели незначимы
- B. Значительное изменение оценок коэффициентов при добавлении/удалении одного из наблюдений
- C. Значения VIF (variance inflation factor) выше 10
- D. Гетероскедастичность остатков

**Поле для ответа:**

--	--	--

**Задание 12**

**Инструкция:** Прочтайте задание и выберите все правильные ответы

Какие из утверждений верны в контексте применения метода наименьших квадратов (МНК)?

- A. При выполнении стандартных предположений, МНК-оценки являются несмешёнными
- B. Оценки МНК минимизируют сумму квадратов отклонений наблюдаемых значений зависимой переменной от предсказанных
- C. МНК применим только при нормальности ошибок
- D. При выполнении стандартных предположений, МНК-оценки обладают наименьшей дисперсией среди всех других линейных несмешенных оценок.

**Поле для ответа:**

--	--	--

**Задание 13**

**Инструкция:** Прочтайте задание и выберите все правильные ответы

Что может служить причиной несостоительности оценок коэффициентов МНК в регрессионной модели?

- A. Пропущенные объясняющие переменные
- B. Эндогенность объясняющих переменных
- C. Гетероскедастичность
- D. Автокорреляция

**Поле для ответа:**

--	--	--

## **Вопросы на выбор нескольких правильных ответов**

### **Компетенция ОПК-5-Э**

#### **Задания закрытого типа**

#### **Повышенный уровень сложности**

#### **Задание 1**

**Инструкция:** Прочтите задание и выберите все правильные ответы

Какие действия относятся к подготовке данных для регрессионного анализа в программной среде?

- A. Удаление пропусков или их корректная обработка
- B. Преобразование категориальных переменных в фиктивные переменные
- C. Визуализация остатков регрессии
- D. Проверка автокорреляции в остатках регрессии

**Поле для ответа:**

--	--	--

#### **Задание 2**

**Инструкция:** Прочтите задание и выберите все правильные ответы

После оценки модели линейной регрессии в статистическом пакете вы видите в итоговой таблице столбец « $\text{Pr}(>|t|)$ ». Что он содержит?

- A.  $t$ -статистики
- B. Вероятность того, что остатки нормально распределены
- C.  $p$ -значения
- D. Стандартные ошибки коэффициентов

**Поле для ответа:**

--	--	--

#### **Задание 3**

**Инструкция:** Прочтите задание и выберите все правильные ответы

Какие из перечисленных ниже программных продуктов часто используются для оценки множественной регрессии?

- A. R
- B. Stata
- C. Microsoft Word
- D. Python (с библиотеками, например statsmodels)

**Поле для ответа:**

--	--	--

#### **Задание 4**

**Инструкция:** Прочтите задание и выберите все правильные ответы

Для автоматизации каких из следующих этапов анализа данных предназначены статистические пакеты?

- A. Импорт данных
- B. Построение регрессионной модели
- C. Написание пояснительной записи в Word
- D. Проверка гетероскедастичности остатков модели

**Поле для ответа:**

--	--	--

**Задание 5**

**Инструкция:** Прочтите задание и выберите все правильные ответы  
Что является типичными результатами неправильной обработки пропущенных значений в данных?

- A. Смещение оценок коэффициентов модели
- B. Потеря репрезентативности выборки
- C. Увеличение точности прогноза модели
- D. Потеря наблюдений

**Поле для ответа:**

--	--	--

**Задание 6**

**Инструкция:** Прочтите задание и выберите все правильные ответы  
В результате оценки регрессионной модели вы хотите проверить отсутствие автокорреляции в остатках. Какой из методов диагностики, реализованных в большинстве статистических пакетов, следует использовать?

- A. Гистограмма остатков
- B. Тест Дарбина–Уотсона
- C. Тест Уайта
- D. F-тест

**Поле для ответа:**

--	--	--

**Задание 7**

**Инструкция:** Прочтите задание и выберите все правильные ответы  
Какие технологии можно использовать для визуализации результатов регрессионного анализа?

- A. ggplot2 (в R)
- B. matplotlib (в Python)
- C. Microsoft Word
- D. Adobe Photoshop

**Поле для ответа:**

--	--	--

### **Задание 8**

**Инструкция:** Прочтайте задание и выберите все правильные ответы  
Какие форматы данных удобны для импорта в статистические пакеты?

- A. csv
- B. xlsx
- C. docx
- D. exe

**Поле для ответа:**

--	--	--

### **Задание 9**

**Инструкция:** Прочтайте задание и выберите все правильные ответы  
В результате оценки регрессионной модели в статистическом пакете среди результатов вы видите столбец с обозначением «s.e.» или «Std.Error.» Что он содержит?

- A. Дисперсию зависимой переменной
- B. Среднюю ошибку прогноза
- C. Стандартную ошибку оценки коэффициента
- D. Среднюю абсолютную ошибку модели

**Поле для ответа:**

--	--	--

### **Задание 10**

**Инструкция:** Прочтайте задание и выберите все правильные ответы  
Какие преимущества даёт использование интегрированной среды разработки (IDE), например, RStudio, JupyterLab, при анализе данных?

- A. Удобное редактирование кода
- B. Интерактивный вывод результатов
- C. Автоматическая генерация научной статьи
- D. Встроенная поддержка графиков и отладки

**Поле для ответа:**

--	--	--

### **Задание 11**

**Инструкция:** Прочтайте задание и выберите все правильные ответы  
После оценки модели в статистическом пакете вы получили, что коэффициент при фиктивной переменной равен  $-5$  и значим на уровне  $1\%$ . В модели есть константа. Какова интерпретация этого результата?

- A. Константа в уравнениях для двух групп (базовой и той, которую кодирует фиктивная переменная) статистически различны
- B. Разницы между группами нет
- C. Группа, обозначенная фиктивной переменной, в среднем имеет значение зависимой переменной на 5 единиц меньше (чем базовая группа)
- D. Все значения переменной равны 5

**Поле для ответа:**

--	--	--

**Задание 12**

**Инструкция:** Прочтайте задание и выберите все правильные ответы  
Какие этапы регрессионного анализа можно выполнять с использованием специализированных библиотек в Python (pandas, statsmodels, scikit-learn)?

- A. Очистка и трансформация данных
- B. Построение и оценка модели
- C. Проверка экономической интерпретируемости модели
- D. Тестирование гипотез

**Поле для ответа:**

--	--	--

**Задание 13**

**Инструкция:** Прочтайте задание и выберите все правильные ответы  
В каких случаях использование программных средств наиболее критично при анализе экономических данных?

- A. При обработке данных объёмом более миллиона наблюдений
- B. При необходимости повторного запуска модели с новыми данными
- C. Для описания структуры данных в Microsoft Word
- D. При использовании сложных моделей (например, с перекрёстными взаимодействиями и множественными регрессорами)

**Поле для ответа:**

--	--	--

**Компетенция ОПК-2-Э**

**Комбинированные задания**

**Повышенный уровень сложности**

**Задание 1**

**Инструкция:** Прочтайте задание и выполните расчеты

Была оценена линейная регрессия  $\hat{y} = 3 + 1.5x$ . найдите прогнозное значение переменной  $y$  в точке  $x = 4$ .

**Поле для ответа:** \_\_\_\_\_

**Задание 2**

**Инструкция:** Прочтайте задание и выполните расчеты

По выборке объема  $n = 20$  была оценена линейная регрессионная модель с  $k = 2$  регрессорами (помимо свободного члена). Коэффициент детерминации оказался равным  $R^2 = 0.81$ . Найдите скорректированный коэффициент детерминации.

**Поле для ответа:** \_\_\_\_\_

### Задание 3

**Инструкция:** Прочтайте задание и выполните расчеты

Выборочный коэффициент корреляции между переменными  $y$  и  $x$  равен  $\text{corr}[y, x] = -0.9$ . Найдите коэффициент детерминации в парной линейной регрессии (с константной) между переменными  $y$  и  $x$ .

**Поле для ответа:** \_\_\_\_\_

### Задание 4

**Инструкция:** Прочтайте задание и выполните расчеты

Коэффициент при переменной  $x$  в уравнении регрессии равен 2.1, стандартная ошибка — 0.7. Рассчитайте  $t$ -статистику для данного коэффициента.

**Поле для ответа:** \_\_\_\_\_

### Задание 5

**Инструкция:** Прочтайте задание и выполните расчеты

Найдите оценку коэффициента  $\beta_1$  в модели парной линейной регрессии  $y = \beta_0 + \beta_1 x + \varepsilon$ , если известно что  $\sum(x_i - \bar{x})^2 = 100$ ,  $\sum(x_i - \bar{x})(y_i - \bar{y}) = 250$ , где  $\bar{x}$  и  $\bar{y}$  – выборочные средние соответствующих переменных.

**Поле для ответа:** \_\_\_\_\_

### Задание 6

**Инструкция:** Прочтайте задание и выполните расчеты

Сумма квадратов остатков в модели линейной регрессии (с константой)  $\text{ESS} = 180$ , а общая сумма квадратов отклонений зависимой переменной от среднего значения  $\text{TSS} = 300$ . Найдите коэффициент детерминации  $R^2$ .

**Поле для ответа:** \_\_\_\_\_

### Задание 7

**Инструкция:** Прочтайте задание и выполните расчеты

В модели  $y_i = \beta_0 + \beta_1 x_i + u_i$  получены:  $\hat{\beta}_1 = 1.8$ ,  $SE = 0.6$ , где  $SE$  – стандартная ошибка коэффициента. Вычислите  $t$ -статистику и определите значим ли коэффициент на уровне 5%. Используйте двусторонний тест и нормальное приближение.

**Поле для ответа:** \_\_\_\_\_

### Задание 8

**Инструкция:** Прочтайте задание и выполните расчеты

В модели линейной регрессии  $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$  при увеличении регрессора  $x_1$  на 1 единицу, зависимая переменная  $y$  в среднем увеличивается на 3.4, при прочих равных условиях. Какова оценка коэффициента регрессии  $\beta_1$ ?

**Поле для ответа:** \_\_\_\_\_

### Задание 9

**Инструкция:** Прочтите задание и выполните расчеты

Было оценено уравнение линейной регрессии  $\hat{y} = 3 + 1.5x$ . Найдите значение остатка (т.е. разницу между фактом и прогнозом) в точке  $x = 4$ , если фактическое значение зависимой переменной в точке  $x = 4$  равно 10.

**Поле для ответа:** \_\_\_\_\_

### Задание 10

**Инструкция:** Прочтите задание и выполните расчеты

Найдите общую сумму квадратов отклонений зависимой переменной от среднего значения (TSS), если сумма квадратов остатков в модели линейной регрессии (ESS) равна 200, а коэффициент детерминации ( $R^2$ ) равен 0.6.

**Поле для ответа:** \_\_\_\_\_

### Задание 11

**Инструкция:** Прочтите задание и выполните расчеты

Какие из приведенных ниже регрессионных моделей линейны по параметрам?

Прокомментируйте ваш ответ.

- A.  $Y = \alpha + \beta_1 X + \beta_2 Z + \varepsilon$ ,
- B.  $Y = \alpha + \beta_1 X - \beta_1 Z + \varepsilon$ ,
- C.  $Y = \alpha + \beta_1 X + \frac{1}{\beta_2} Z + \varepsilon$ ,
- D.  $Y = \alpha + \beta_1 \log(X) + \beta_2 \log(Z) + \varepsilon$ ,
- E.  $\log(Y) = \alpha + \beta_1 X + \beta_2 Z + \varepsilon$ ,
- F.  $Y = \alpha + \beta_1 X + \beta_2 \frac{1}{Z} + \varepsilon$ .

**Поле для ответа:** \_\_\_\_\_

### Задание 12

**Инструкция:** Прочтите задание и выполните расчеты

В модели линейной регрессии с двумя объясняющими переменными ( $k$ ) и свободным членом, оцененной по выборке из  $n = 25$  наблюдений, получена сумма квадратов остатков  $ESS = 120$ . Найдите МНК-оценку для дисперсии ошибок в регрессии.

**Поле для ответа:** \_\_\_\_\_

### Задание 13

**Инструкция:** Прочтите задание и выполните расчеты

При построении модели линейной регрессии  $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$  получено:  $\hat{\beta}_1 = 1.8$ , стандартная ошибка коэффициента  $\hat{\beta}_1$  равна  $SE = 0.6$ . Найдите 95% доверительный интервал для коэффициента  $\beta_1$ , если 2.5%-я точка распределения Стьюдента с соответствующими степенями свободы равна  $t = 2.1$ .

**Поле для ответа:** \_\_\_\_\_

#### Задание 14

**Инструкция:** Прочтайте задание и выполните расчеты

Пусть  $\hat{\beta}_0$  и  $\hat{\beta}_1$  – оценки МНК параметров  $\beta_0$  и  $\beta_1$  в линейной регрессии

$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ ,  $i = 1, \dots, n$ . Чему равна сумма остатков в модели,

$$\sum_{i=1}^n (Y_i - \hat{\beta}_0 + \hat{\beta}_1 X_i) ?$$
 Ответ обоснуйте.

**Поле для ответа:** \_\_\_\_\_

#### Задание 15

**Инструкция:** Прочтайте задание и выполните расчеты

Рассмотрим линейную регрессию  $Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i$ , где  $Y$  – объем продаж товара (руб.),  $X$  – цена товара (руб.). Что и как изменится в модели, если зависимая переменная ( $Y$ ) будет измеряться не в рублях, а в тыс. рублей?

- A. оценка свободного члена
- B. оценка коэффициента наклона
- C. общая сумма квадратов
- D. оценка дисперсии ошибок
- E. коэффициент детерминации

**Поле для ответа:** \_\_\_\_\_

#### Задание 16

**Инструкция:** Прочтайте задание и выполните расчеты

В исследовании была построена регрессионная модель  $\hat{y} = 5 + 0.8x_1 - 0.4x_2$ . Найдите прогнозное значение зависимой переменной  $y$  в точке  $x_1 = 10$ ,  $x_2 = 5$ .

**Поле для ответа:** \_\_\_\_\_

#### Задание 17

**Инструкция:** Прочтайте задание и выполните расчеты

Дана выборка значений объясняющей переменной  $x$ : [2, 4, 6, 8, 10] и зависимой переменной  $y$ : [3, 5, 7, 9, 11]. Найдите МНК-оценки коэффициентов  $a$  и  $b$  в линейной регрессии  $y = a + bx + \varepsilon$ .

**Поле для ответа:** \_\_\_\_\_

**Компетенция ОПК-5-Э**  
**Комбинированные задания**  
**Повышенный уровень сложности**

**Задание 1**

**Инструкция:** Прочтите задание и выполните расчеты

По результатам регрессионного анализа в статистическом пакете найдены:

- Общая сумма квадратов отклонений зависимой переменной от среднего значения (TSS) = 500
- Сумма квадратов отклонений, объясненная регрессией (RSS) = 300

Найдите сумму квадратов остатков в регрессии (ESS).

**Поле для ответа:** \_\_\_\_\_

**Задание 2**

**Инструкция:** Прочтите задание и выполните расчеты

По 100 наблюдениям оценивается линейная регрессия с 2 регрессорами (помимо свободного члена). Статистический пакет выдал оценку коэффициента при одном из регрессоров (= 2.4) и стандартную ошибку данной оценки (= 0.4). Проверьте на 5% уровне значимости гипотезу о том, что данный регрессор незначимо отличен от 0, если известно, что 2.5%-я точка распределения Стьюдента с 97 степенями свободы равна 1.98.

**Поле для ответа:** \_\_\_\_\_

**Задание 3**

**Инструкция:** Прочтите задание и выполните расчеты

Вы оценили следующую модель:

$$health\_ex_i = \beta_0 + \beta_1 \cdot \log(income) + \beta_2 \cdot age + u$$

где

*health\_ex* – индивидуальные ежемесячные расходы на здоровье (в рублях),

*income* – доход (в рублях),

*age* – возраст (в годах).

По результатам оценки модели в статистическом пакете получено:

$$\hat{\beta}_1 = 0.8$$

$$\hat{\beta}_2 = 0.04$$

Как правильно интерпретировать эти коэффициенты?

**Поле для ответа:** \_\_\_\_\_

**Задание 4**

**Инструкция:** Прочтите задание и выполните расчеты

Вы оценили в статистическом пакете модель линейной регрессии. После добавления в модель новой объясняющей переменной значение коэффициента детерминации  $R^2$  уменьшилось. Что можно сказать о такой ситуации? Выберите один вариант из предложенных. Обоснуйте ответ.

А. Это возможно, если новая переменная незначима

- B. Это возможно, если новая переменная коррелирует с ошибкой  
C. Это возможно, если новая переменная коррелирует с зависимой переменной  
D. Это невозможно

**Поле для ответа:**

--	--	--

**Обоснование:** \_\_\_\_\_.

### Задание 5

**Инструкция:** Прочтите задание и выполните расчеты

Оценивается линейная регрессия. В статистическом пакете найдены следующие статистики:

- Residual standard error ( $\sigma$ ) = 5
- Степени свободы ( $df$ ) = 20

Найдите сумму квадратов остатков в регрессии ( $ESS$ ).

**Поле для ответа:** \_\_\_\_\_

### Задание 6

**Инструкция:** Прочтите задание и выполните расчеты

В статистическом пакете оценена линейная регрессионная модель  $\hat{y} = 2 + 0.5x_1 + 1.1x_2$ .

Найдите прогнозное значение зависимой переменной  $y$  в точке  $x_1 = 4, x_2 = 6$ .

**Поле для ответа:** \_\_\_\_\_

### Задание 7

**Инструкция:** Прочтите задание и выполните расчеты

После устранения гетероскедастичности методами робастных ошибок в статистическом пакете, стандартная ошибка оценки коэффициента уменьшилась с 0.6 до 0.4. Во сколько раз выросла t-статистика для данного коэффициента?

**Поле для ответа:** \_\_\_\_\_

### Задание 8

**Инструкция:** Прочтите задание и выполните расчеты

В результате оценки регрессионной модели, включающей регрессор  $x_1$ , в статистическом пакете вы получили следующий результат:

Переменная Оценка коэффициента Стандартная ошибка

$x_1$	2.4	0.8
-------	-----	-----

Постройте 95% доверительный интервал для коэффициента при переменной  $x_1$ , если известно, что квантиль уровня 0.975 распределения Стьюдента с соответствующими степенями свободы равен 1.96.

**Поле для ответа:** \_\_\_\_\_

### **Задание 9**

**Инструкция:** Прочтите задание и выполните расчеты

Пусть  $Q_1, \dots, Q_4$  – фиктивные переменные для сезонов (например, переменная  $Q_3$  равна 1, если наблюдение произведено осенью и 0, иначе). Почему в линейную регрессию с константой включается лишь три из этих фиктивных переменных. Что будет если включить все четыре?

**Поле для ответа:** \_\_\_\_\_

### **Задание 10**

**Инструкция:** Прочтите задание и выполните расчеты

В результате оценки модели линейной регрессии  $y = \alpha + \beta x + \varepsilon$  в статистическом пакете найдены  $\hat{\beta} = 0.8$  (оценка коэффициента  $\beta$ ),  $\bar{x} = 5$  (выборочное среднее регрессора) и  $\bar{y} = 7$  (выборочное среднее зависимой переменной). Найдите МНК-оценку  $\hat{\alpha}$  свободного члена  $\alpha$ .

**Поле для ответа:** \_\_\_\_\_

### **Задание 11**

**Инструкция:** Прочтите задание и выполните расчеты

В статье приводится результат оценки линейной регрессии  $\hat{y} = 1.1_{(2.37)} - 0.7_{(-0.4)} x$ . Как вы

думаете, что именно указано в скобках? Ответ обоснуйте. Варианты ответа:

- A.  $p$ -значения
- B.  $t$ -статистики
- C. Стандартные ошибки коэффициентов
- D. Показатели VIF для каждого коэффициента

**Поле для ответа:** \_\_\_\_\_

### **Задание 12**

**Инструкция:** Прочтите задание и выполните расчеты

В статистическом пакете оценена лог-линейная модель  $\ln(y) = 1 + 2 \cdot \ln(x) + \varepsilon$ . На сколько процентов в среднем изменится зависимая переменная  $y$  при увеличении  $x$  на 1%?

**Поле для ответа:** \_\_\_\_\_

### **Задание 13**

**Инструкция:** Прочтите задание и выполните расчеты

В статистическом пакете найдены остатки в парной линейной регрессии со свободным членом:  $e=[1.5, -0.5, 2.0, -1.0, 1.0]$ . Найдите МНК-оценку дисперсии ошибок.

**Поле для ответа:** \_\_\_\_\_

### Задание 14

**Инструкция:** Прочтайте задание и выполните расчеты

В статистическом пакете была оценена линейная регрессия  $y = \beta_0 + \beta_1 x + \varepsilon$ .

Коэффициент  $\beta_1$  оказался незначимым на уровне  $\alpha$ . Из этого следует, что доверительный интервал для коэффициента  $\beta_1$  уровня  $1 - \alpha$  (выберите правильный вариант ответа и прокомментируйте его):

- A. содержит 0;
- B. содержит  $\alpha$ ;
- C. содержит  $1 - \alpha$ ;
- D. не содержит истинное значение  $\beta_1$ .

**Поле для ответа:** \_\_\_\_\_

### Задание 15

**Инструкция:** Прочтайте задание и выполните расчеты

Оцененное уравнение линейной регрессии имеет вид:  $\hat{y} = 10.3 + 15.6x_1 + 0.1 \cdot 10^{-6}x_2$ .

Можно ли без дополнительной информации утверждать, что коэффициент при регрессоре  $x_2$  не значим? Ответ обоснуйте.

**Поле для ответа:** \_\_\_\_\_

### Задание 16

**Инструкция:** Прочтайте задание и выполните расчеты

Пусть UR – уровень безработицы (%), DUR – длительность выплаты пособия (года).

Результат оценки линейной регрессии в статистическом пакете приведен в распечатке ниже.

The screenshot shows the gretl software window titled "gretl: модель 2". The menu bar includes "Файл", "Правка", "Тесты", "Сохранить", "Графики", "Анализ", and "LaTeX". The main text area displays the following regression output:

Модель 2: МНК, использованы наблюдения 1-20  
Зависимая переменная: UR

	Коэффициент	Ст. ошибка	t-статистика	P-значение	
const	4,33174	1,59142	2,722	0,0140	**
DUR	1,61011	0,596035	2,701	0,0146	**
Среднее зав. перемен	8,035000	Ст. откл. зав. перемен	4,170987		
Сумма кв. остатков	235,1945	Ст. ошибка модели	3,614742		
R-квадрат	0,288465	Испр. R-квадрат	0,248936		
F(1, 18)	7,297437	P-значение (F)	0,014609		
Лог. правдоподобие	-53,02558	Крит. Акаике	110,0512		
Крит. Шварца	112,0426	Крит. Хеннана-Куинна	110,4399		

Проверьте на 5%-м уровне значимости гипотезу о том, что коэффициент при DUR равен 2. Воспользуйтесь тем, что 2.5%-я точка распределения Стьюдента с 18 степенями свободы равна 2.1.

**Поле для ответа:** \_\_\_\_\_

### Задание 17

**Инструкция:** Прочтите задание и выполните расчеты

Пусть UR – уровень безработицы (%), DUR – длительность выплаты пособия (годы). Результат оценки линейной регрессии в статистическом пакете приведен в распечатке ниже.

The screenshot shows the gretl software window titled "gretl: модель 2". The menu bar includes Файл, Правка, Тесты, Сохранить, Графики, Анализ, LaTeX. The main text area displays the following information:

Модель 2: МНК, использованы наблюдения 1-20  
Зависимая переменная: UR

	Коэффициент	Ст. ошибка	t-статистика	P-значение	
const	4,33174	1,59142	2,722	0,0140	**
DUR	1,61011	0,596035	2,701	0,0146	**
Среднее зав. перемен	8,035000	Ст. откл. зав. перемен	4,170987		
Сумма кв. остатков	235,1945	Ст. ошибка модели	3,614742		
R-квадрат	0,288465	Испр. R-квадрат	0,248936		
F(1, 18)	7,297437	P-значение (F)	0,014609		
Лог. правдоподобие	-53,02558	Крит. Акаике	110,0512		
Крит. Шварца	112,0426	Крит. Хеннана-Куинна	110,4399		

Выпишите оцененное уравнение и оцените его качество с помощью теста Фишера. Используйте уровень значимости 0.05.

**Поле для ответа:** \_\_\_\_\_

### Компетенция ОПК-2-Э

**Задания открытого типа**

**Высокий уровень сложности**

### Задание 1

**Инструкция:** Прочтите задание и дайте развернутый ответ

Каковы последствия эндогенности? Объясните зачем и в каких случаях для оценки регрессионной модели используются инструментальные переменные.

**Поле для ответа:** \_\_\_\_\_

### Задание 2

**Инструкция:** Прочтите задание и дайте развернутый ответ

Опишите алгоритм оценки модели линейной регрессии двухшаговым методом наименьших квадратов (2SLS).

**Поле для ответа:** \_\_\_\_\_

### Задание 3

**Инструкция:** Прочтите задание и дайте развернутый ответ

Расскажите про гетероскедастичность, её признаки, последствия и способы корректировки.

**Поле для ответа:** \_\_\_\_\_

#### **Задание 4**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Что означает автокорреляция остатков в регрессионной модели и чем она опасна? Кратко перечислите способы её обнаружения и устранения.

**Поле для ответа:** \_\_\_\_\_

#### **Задание 5**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Почему важно проверять правильность спецификации регрессионной модели? Какие типичные ошибки спецификации могут возникать и как их диагностировать?

**Поле для ответа:** \_\_\_\_\_

#### **Задание 6**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Объясните различие между  $R^2$  (коэффициент детерминации) и скорректированным  $R^2$ .

Почему нельзя полагаться на  $R^2$  при сравнении вложенных моделей?

**Поле для ответа:** \_\_\_\_\_

#### **Задание 7**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Найдите оценку МНК параметра  $\alpha$  в регрессии  $y_i = \alpha + \varepsilon_i, i = 1, \dots, n$ .

**Поле для ответа:** \_\_\_\_\_

#### **Задание 8**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Найдите оценку МНК параметра  $\beta$  в регрессии  $y_i = \beta x_i + \varepsilon_i, i = 1, \dots, n$ .

**Поле для ответа:** \_\_\_\_\_

#### **Задание 9**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Какие ключевые предпосылки лежат в основе метода наименьших квадратов (МНК)? Что происходит с оценками МНК, если хотя бы одна из них нарушается?

**Поле для ответа:** \_\_\_\_\_

#### **Задание 10**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Модель линейной регрессии имеет коэффициент детерминации  $R^2 = 0.95$ , но один из коэффициентов (при переменной  $X_1$ ) имеет  $p$ -значение = 0.55. Можно ли утверждать, что переменная  $X_1$  важна для модели?

**Поле для ответа:** \_\_\_\_\_

### **Задание 11**

**Инструкция:** Прочтите задание и дайте развернутый ответ

Можно ли на основе значимости коэффициента в линейной регрессии утверждать о причинной связи между соответствующим регрессором и зависимой переменной?

**Поле для ответа:** \_\_\_\_\_

### **Задание 12**

**Инструкция:** Прочтите задание и дайте развернутый ответ

Некто оценил регрессию по 300 наблюдениям и получил  $\hat{Y} = -20 + 10X, \bar{Y} = 200$ . После этого он добавил одно наблюдение, переоценил регрессию, и оказалось, что TSS (общая сумма квадратов отклонений зависимой переменной от среднего значения) и RegSS (сумма квадратов отклонений, объясненная регрессией) не изменились. Можете ли вы сказать, какое наблюдение было добавлено?

**Поле для ответа:** \_\_\_\_\_

### **Задание 13**

**Инструкция:** Прочтите задание и дайте развернутый ответ

Почему важно исследовать остатки регрессионной модели? Что можно понять из графика остатков и как он помогает улучшить модель?

**Поле для ответа:** \_\_\_\_\_

### **Задание 14**

**Инструкция:** Прочтите задание и дайте развернутый ответ

Что означает наличие произведения непрерывной и фиктивной переменных в регрессии? Приведите пример и объясните интерпретацию коэффициентов.

**Поле для ответа:** \_\_\_\_\_

### **Задание 15**

**Инструкция:** Прочтите задание и дайте развернутый ответ

Объясните, в чём заключается проблема эндогенности в регрессионном анализе. Как её можно обнаружить и устранить?

**Поле для ответа:** \_\_\_\_\_

### **Задание 16**

**Инструкция:** Прочтите задание и дайте развернутый ответ

Какие признаки указывают на наличие гетероскедастичности в регрессионной модели? Каковы её последствия и способы коррекции?

**Поле для ответа:** \_\_\_\_\_

### **Задание 17**

**Инструкция:** Прочтите задание и дайте развернутый ответ

Ваш коллега утверждает, что модель с  $R^2 = 0.9$  всегда лучше модели с  $R^2 = 0.7$  пусть даже и с меньшим набором регрессоров. Когда модель с более низким  $R^2$  может быть предпочтительнее?

**Поле для ответа:** \_\_\_\_\_

### **Компетенция ОПК-5-Э**

#### **Задания открытого типа**

#### **Высокий уровень сложности**

##### **Задание 1**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

В статистическом пакете вы оценили две модели зависимости потребительских расходов от дохода и других факторов. Программа вывела следующую информацию:

Модель	$R^2$	AIC	BIC	Число параметров
--------	-------	-----	-----	------------------

A	0.76	220.4	230.1	6
---	------	-------	-------	---

B	0.76	215.2	223.5	4
---	------	-------	-------	---

Какую модель следует предпочесть? Ответ поясните.

**Поле для ответа:** \_\_\_\_\_

##### **Задание 2**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Вы оцениваете регрессионную модель в статистическом пакете. Как проверить значимость коэффициентов?

**Поле для ответа:** \_\_\_\_\_

##### **Задание 3**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Вы оценили в статистическом пакете регрессионную модель, однако в остатках обнаружилась гетероскедастичность. Объясните, почему гетероскедастичность нарушает корректность выводов и какие подходы позволяют получить корректные стандартные ошибки.

**Поле для ответа:** \_\_\_\_\_

##### **Задание 4**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Вы оцениваете модель влияния образования (educ) на доход (income). Также в модель включена переменная age (возраст). Поскольку educ может быть эндогенной, вы применяете метод инструментальных переменных: используете переменную distance (расстояние до университета) как инструмент для educ.

После первого шага 2SLS (регрессии educ на distance и age) статистический пакет вывел:

- Коэффициент при distance: -0.05
- t-статистика для коэффициента при distance: -1.21
- p-значение для коэффициента при distance: 0.23

- $R^2$  регрессии из 1 шага: 0.02
- F-статистика регрессии из 1 шага: 1.5

Что можно сказать об инструменте `distance`? Как это влияет на вашу оценку эффекта образования?

**Поле для ответа:** \_\_\_\_\_

### Задание 5

**Инструкция:** Прочтите задание и дайте развернутый ответ

Как можно использовать возможности программной среды R для визуализации остатков регрессионной модели? Что можно визуально оценить с помощью данных графиков?

**Поле для ответа:** \_\_\_\_\_

### Задание 6

**Инструкция:** Прочтите задание и дайте развернутый ответ

Как можно проверить наличие автокорреляции в регрессионной модели в пакете R? Какие выводы можно сделать из анализа автокорреляции?

**Поле для ответа:** \_\_\_\_\_

### Задание 7

**Инструкция:** Прочтите задание и дайте развернутый ответ

Какие критерии качества оценки регрессионной модели можно получить в пакете R и как интерпретировать их значения?

**Поле для ответа:** \_\_\_\_\_

### Задание 8

**Инструкция:** Прочтите задание и дайте развернутый ответ

Как реализовать проверку гетероскедастичности в регрессионной модели в пакете R?

Какие меры могут быть приняты при её наличии?

**Поле для ответа:** \_\_\_\_\_

### Задание 9

**Инструкция:** Прочтите задание и дайте развернутый ответ

Менеджер кабельной телевизионной сети планирует изменить одноразовую плату за установку оборудования для просмотра кабельного телевидения (`install_fee`, % от среднего душевого дохода) для того, чтобы увеличить долю квартир, имеющих подписку на кабельное телевидение (`cable_share`, %). Он предполагает, что стоит также учитывать число транслируемых платных кабельных каналов (`cable_channels`), число доступных бесплатных эфирных каналов (`air_channels`), ежемесячную плату за кабельное телевидение (`monthly_fee`, % от среднего душевого дохода) и сколько лет функционирует сеть (`age`, годы).

Менеджер сформулировал несколько интересующих его вопросов, которые он собирается проверить с помощью регрессионного уравнения

$$\text{cable\_share} = \beta_1 + \beta_2 * \text{cable\_channels} + \beta_3 * \text{age} + \beta_4 * \text{air\_channels} + \beta_5 * \text{monthly\_fee} + \beta_6 * \text{install\_fee} + \varepsilon$$

Запишите следующие гипотезы на языке коэффициентов:

1. Не удалось обнаружить факторы, влияющие на долю квартир, имеющих подписку на кабельное телевидение.

2. Одноразовая плата за установку оборудования для просмотра кабельного телевидения не влияет на число квартир, имеющих подписку на кабельное телевидение.

3. Каждый новый бесплатный эфирный канал при прочих равных условиях уменьшает долю квартир, имеющих подписку на кабельное телевидение, в среднем при прочих равных условиях на 4 п.п.

4. Число кабельных каналов, ежемесячная плата за кабельное телевидение и одноразовая плата за установку оборудования для просмотра кабельного телевидения не влияют на долю квартир, имеющих подписку на кабельное телевидение.

**Поле для ответа:** \_\_\_\_\_

### Задание 10

**Инструкция:** Прочтите задание и дайте развернутый ответ

Менеджер кабельной телевизионной сети планирует изменить одноразовую плату за установку оборудования для просмотра кабельного телевидения (*install\_fee*, % от среднего душевого дохода) для того, чтобы увеличить долю квартир, имеющих подписку на кабельное телевидение (*cable\_share*, %). Он предполагает, что стоит также учитывать число транслируемых платных кабельных каналов (*cable\_channels*), число доступных бесплатных эфирных каналов (*air\_channels*), ежемесячную плату за кабельное телевидение (*monthly\_fee*, % от среднего душевого дохода) и сколько лет функционирует сеть (*age*, годы).

Менеджер сформулировал несколько интересующих его вопросов, которые он собирается проверить с помощью регрессионного уравнения

$$cable\_share = \beta_1 + \beta_2 * cable\_channels + \beta_3 * age + \beta_4 * air\_channels + \beta_5 * monthly\_fee + \beta_6 * install\_fee + \varepsilon$$

Используя распечатки ниже, проверьте следующие гипотезы:

1. Не удалось обнаружить факторы, влияющие на долю квартир, имеющих подписку на кабельное телевидение.

2. Одноразовая плата за установку оборудования для просмотра кабельного телевидения не влияет на число квартир, имеющих подписку на кабельное телевидение.

3. Число кабельных каналов, ежемесячная плата за кабельное телевидение и одноразовая плата за установку оборудования для просмотра кабельного телевидения не влияют на долю квартир, имеющих подписку на кабельное телевидение.

Модель 1: МНК, использованы наблюдения 1-40

Зависимая переменная: *cable\_share*

Робастные оценки стандартных ошибок (с поправкой на гетероскедастичность), вариант HC1

	Коэффициент <i>t</i>	Ст. ошибка	<i>t</i> -статистика	<i>p</i> -значение	
const	72,2250	21,4326	3,370	0,0019	***
cable_channels	0,842544	0,782758	1,076	0,2893	
age	2,12862	0,637828	3,337	0,0021	***
air_channels	-4,34894	1,13575	-3,829	0,0005	***
monthly_fee	-301,425	206,548	-1,459	0,1536	
install_fee	-17,5727	46,1525	-0,3808	0,7058	

Среднее завис. перемен	44,57971	Ст. откл. завис. перемен	18,47249
Сумма кв. остатков	6605,832	Ст. ошибка модели	13,93877
R-квадрат	0,503623	Исправ. R-квадрат	0,430626
F(5, 34)	9,375521	P-значение (F)	0,000011
Лог. правдоподобие	-158,8941	Крит. Акаике	329,7882

Крит. Шварца

339,9215

Крит. Хеннана-Куинна

333,4521

Модель 2: МНК, использованы наблюдения 1-40

Зависимая переменная: cable\_share

Робастные оценки стандартных ошибок (с поправкой на гетероскедастичность), вариант HC1

	<i>Коэффициен т</i>	<i>Ст. ошибка</i>	<i>t- статистика</i>	<i>p-значение</i>
const	47,5765	8,68743	5,476	<0,0001 ***
age	2,32222	0,582675	3,985	0,0003 ***
air_channels	-3,59068	0,984869	-3,646	0,0008 ***
Среднее завис. перемен	44,57971	Ст. откл. завис. перв		18,47249
Сумма кв. остатков	7357,165	Ст. ошибка модели		14,10115
R-квадрат	0,447166	Исправ. R-квадрат		0,417283
F(2, 37)	20,30335	P-значение (F)		1,12e-06
Лог. правдоподобие	-161,0486	Крит. Акаике		328,0971
Крит. Шварца	333,1637	Крит. Хеннана-Куинна		329,9290

Чётко выпишите проверяемые гипотезы, аргументируйте ответы.

Поле для ответа: \_\_\_\_\_

**Задание 11****Инструкция:** Прочитайте задание и дайте развернутый ответ

Менеджер кабельной телевизионной сети планирует изменить одноразовую плату за установку оборудования для просмотра кабельного телевидения (*install\_fee*, % от среднего душевого дохода) для того, чтобы увеличить долю квартир, имеющих подписку на кабельное телевидение (*cable\_share*, %). Он предполагает, что стоит также учитывать число транслируемых платных кабельных каналов (*cable\_channels*), число доступных бесплатных эфирных каналов (*air\_channels*), ежемесячную плату за кабельное телевидение (*monthly\_fee*, % от среднего душевого дохода) и сколько лет функционирует сеть (*age*, годы).

Менеджер сформулировал несколько интересующих его вопросов, которые он собирается проверить с помощью регрессионного уравнения

$$\text{cable\_share} = \beta_1 + \beta_2 * \text{cable\_channels} + \beta_3 * \text{age} + \beta_4 * \text{air\_channels} + \beta_5 * \text{monthly\_fee} + \beta_6 * \text{install\_fee} + \varepsilon$$

Результаты оценки модели приведены ниже. Регрессоры *cable\_channels*, *monthly\_fee* и *install\_fee* оказались незначимы и были исключены из модели.

Модель 1: МНК, использованы наблюдения 1-40

Зависимая переменная: cable\_share

Робастные оценки стандартных ошибок (с поправкой на гетероскедастичность), вариант HC1

	<i>Коэффициен т</i>	<i>Ст. ошибка</i>	<i>t- статистика</i>	<i>p-значение</i>
const	47,5765	8,68743	5,476	<0,0001 ***
age	2,32222	0,582675	3,985	0,0003 ***
air_channels	-3,59068	0,984869	-3,646	0,0008 ***
Среднее завис. перемен	44,57971	Ст. откл. завис. перв		18,47249

Сумма кв. остатков	7357,165	Ст. ошибка модели	14,10115
R-квадрат	0,447166	Исправ. R-квадрат	0,417283
F(2, 37)	20,30335	P-значение (F)	1,12e-06
Лог. правдоподобие	-161,0486	Крит. Акаике	328,0971
Крит. Шварца	333,1637	Крит. Хеннана-Куинна	329,9290

Проверьте значимость данной модели в целом и отдельных коэффициентов, проверьте качество подгонки, интерпретируйте коэффициенты.

### Решение:

Гипотеза о незначимости модели (равно как и отдельных коэффициентов) отвергается на уровне значимости 0.05.

#### Интерпретация коэффициентов:

- каждый дополнительный год функционирования сети в среднем увеличивает долю квартир, имеющих подписку на кабельное телевидение, на 2.3 п.п.
- каждый дополнительный бесплатный канал в среднем снижает долю квартир, имеющих подписку на кабельное телевидение, на 3.6 п.п.

### Задание 12

**Инструкция:** Прочтите задание и дайте развернутый ответ

Вы исследуете как количество сотрудников в фирме (employees) влияет на годовую выручку (revenue). Предполагается, что employees может быть эндогенной переменной (например, фирмы с высокой выручкой нанимают больше людей).

В качестве инструмента используется переменная regional\_wage (средняя региональная зарплата). Также в модель включена переменная sector (отрасль).

После оценки моделей МНК и 2SLS (двуухшаговый МНК) вы провели тест Хаусмана. Статистический пакет привел следующие результаты:

- $\chi^2 = 6.32$  (значение статистики в teste Хаусмана),
- $p$ -значение = 0.012.

Какой вывод позволяет сделать тест Хаусмана? Какая из оценок, оценка МНК или оценка 2SLS, предпочтительна? Используйте уровень значимости 5%.

**Поле для ответа:** \_\_\_\_\_

### Задание 13

**Инструкция:** Прочтите задание и дайте развернутый ответ

В результате оценки двух моделей в статистическом пакете получены следующие суммы квадратов остатков (residual sum of squares, RSS):

- $\text{RSS}_1 = 150$  (базовая модель),
- $\text{RSS}_2 = 120$  (расширенная модель).
- Обе модели оценены по одной и той же выборке, состоящей из  $n = 50$  наблюдений.
- Базовая модель содержит  $k_1 = 2$  коэффициента, расширенная модель —  $k_2 = 4$  коэффициента.

Проверьте по F-критерию, является ли расширенная модель статистически лучше на уровне значимости 5%. Используйте критическое значение  $F_{0.05}(2, 46) = 3.20$ .

**Поле для ответа:** \_\_\_\_\_

### **Задание 14**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Опишите пример сценария, когда необходимо использовать рабочие стандартные ошибки. Как это реализовать в R?

**Поле для ответа:** \_\_\_\_\_

### **Задание 15**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

При выводе статистик, связанных с методом наименьших квадратов, статистические пакеты помимо коэффициента детерминации обычно приводят и скорректированный коэффициент детерминации. Объясните в чём состоит отличие в использовании коэффициента детерминации и скорректированного коэффициента детерминации? В каких случаях эти показатели совпадают?

**Поле для ответа:** \_\_\_\_\_

### **Задание 16**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Большинство статистических пакетов позволяет реализовать метод инструментальных переменных. Расскажите про данный метод.

**Поле для ответа:** \_\_\_\_\_

### **Задание 17**

**Инструкция:** Прочтайте задание и дайте развернутый ответ

Как проверить корректность спецификации модели с использованием статистического пакета R?

**Поле для ответа:** \_\_\_\_\_

## **5 Средства оценки индикаторов достижения компетенций**

Таблица 5

### **Средства оценки индикаторов достижения компетенций**

<b>Коды компетенций</b>	<b>Индикаторы компетенций (в соотв. с Таблицей 1)</b>	<b>Средства оценки (в соотв. с Таблицами 5, 7)</b>
ОПК-2-Э	ИД.ОПК-2.1. ИД.ОПК-2.2 ИД.ОПК-2.3	домашнее задание (ДЗ), тест
ОПК-5-Э	ИД.ОПК-5.1. ИД.ОПК-5.3. ИД.ОПК-5.4. ИД.ОПК-5.5.	домашнее задание (ДЗ), тест

Таблица 6

**Описание средств оценки индикаторов достижения компетенций**

<b>Средства оценки (в соот. с Таблицами 5, 7)</b>	<b>Рекомендованный план выполнения работы</b>
Домашнее задание	<p>Магистрант должен быть готовым в ходе подготовки и представления домашнего задания по темам дисциплины, выполнять следующие действия:</p> <ul style="list-style-type: none"> <li>— Для решения поставленных научно-исследовательских задач различного характера применяет продвинутые инструментальные методы экономического анализа</li> <li>— Использует современные информационные технологии и программные средства при решении профессиональных задач.</li> </ul>
Тест	<p>Магистрант должен быть готовым в ходе подготовки к тесту, выполнять следующие действия:</p> <ul style="list-style-type: none"> <li>— Для решения поставленных научно-исследовательских задач различного характера применяет продвинутые инструментальные методы экономического анализа</li> <li>— Использует современные информационные технологии и программные средства при решении профессиональных задач.</li> </ul>