

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:

ed68fd4b85b778e0f0b1bfea5dbc56cf4148f1225917e739a70e315177f88591

**Автономная некоммерческая образовательная организация высшего образования  
«Европейский университет в Санкт-Петербурге»**

**Школа вычислительных социальных наук**

УТВЕРЖДАЮ:

Ректор

В.В. Волков

« 26 » *августа* 2025 г.

Протокол УС № 2 от 26.02 2025 г.



**Рабочая программа дисциплины  
Панельные, пространственные и дискретные модели**

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:**

Раскина Ю.В., кандидат экономических наук, доцент Школы вычислительных социальных наук АНООВО «ЕУСПб»

**Рецензент:**

Соколов М.В., кандидат экономических наук, старший научный сотрудник по направлению Экономика, Школа вычислительных социальных наук АНООВО «ЕУСПб»

Рабочая программа дисциплины **«Панельные, пространственные и дискретные модели»**, 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 **«Панельные, пространственные и дискретные модели»** является обязательной дисциплиной части, формируемой участниками образовательных отношений, Блока 1 «Дисциплины (модули)» образовательной программы «Данные, знания, экономика, нарративы».

Дисциплина **«Панельные, пространственные и дискретные модели»** направлена на формирование у обучающихся навыков применения статистических методов для анализа данных с повторяющимися наблюдениями, пространственной структурой и ограниченной зависимой переменной. Особое внимание в дисциплине уделяется аналитической работе с данными с использованием современных моделей, включая выбор соответствующей спецификации модели, интерпретацию оценённых коэффициентов, проверку статистических гипотез и выявление устойчивых закономерностей в социально-экономических процессах.

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 .....	7
5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ .....	7
5.1 Содержание дисциплины.....	7
5.2 Структура дисциплины.....	9
6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ.....	9
6.1 Общие положения .....	10
6.2 Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины .....	10
6.3 Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине.....	12
6.4 Перечень литературы для самостоятельной работы обучающегося:.....	13
6.5 Перечень учебно-методического обеспечения для самостоятельной работы.....	13
7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ.....	13
7.1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации .....	13
7.2 Контрольные задания для текущей аттестации.....	15
7.3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации.....	20
7.4 Типовые задания к промежуточной аттестации.....	23
7.5 Средства оценки индикаторов достижения компетенций.....	35
8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА .....	36
8.1. Основная литература.....	36
8.2. Дополнительная литература.....	36
9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА .....	36
9.1 Программное обеспечение .....	37
9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины:.....	37
9.3 Лицензионные электронные ресурсы библиотеки Университета .....	37
9.4 Электронная информационно-образовательная среда Университета.....	38
10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА.....	38
ПРИЛОЖЕНИЕ 1 .....	40

## 1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

**Цель** освоения дисциплины «Панельные, пространственные и дискретные модели» — сформировать у студентов навыки разработки, оценки и интерпретации эконометрических моделей для анализа сложных типов данных, включая панельные, пространственные и дискретные данные.

**Задачи** освоения дисциплины «Панельные, пространственные и дискретные модели» включают:

- освоить принципы построения и интерпретации моделей с фиксированными и случайными эффектами;
- изучить методы оценки динамических панельных моделей;
- овладеть подходами к моделированию пространственной автокорреляции;
- научиться применять модели бинарного, порядкового, мультиномиального и счётного выбора;
- изучить методы анализа длительности состояний;
- развить способности критически оценивать спецификацию модели и качество её прогнозов.

## 2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ

В результате изучения учебной дисциплины обучающийся должен овладеть следующими компетенциями: профессиональными (ПК). 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)
ПК-3 Способен составлять прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом на основании самостоятельно разработанных моделей с использованием математических, статистических и инструментальных методов, а также современных интеллектуальных технологий	ИД.ПК-3.1. Применяет методику прогнозирования основных показателей деятельности предприятия, экономического развития отрасли, региона и экономики в целом ИД.ПК-3.2. Осуществляет микроэкономическое и макроэкономическое моделирование с применением современных инструментов	Знать: основные понятия и категории макроэкономической статистики, ее основные показатели на разных уровнях агрегирования; основные положения теории фирмы, современные подходы к анализу несовершенной конкуренции З (ПК-3)
		Уметь: анализировать показатели результатов финансово-экономической деятельности экономических субъектов разных уровней У (ПК-3)
		Владеть: современными методиками расчета и анализа социально-экономических, финансовых показателей, характеризующих экономические и финансовые процессы и явления на микро- и макроуровне; методами и приемами анализа экономических и финансовых процессов В (ПК-3)
ПК-4 Способен работать с различными источниками статистической информации, разрабатывать и совершенствовать	ИД.ПК-4.1. Работает с различными источниками статистической информации ИД.ПК-4.2. Производит экспериментальные статистические расчеты с	Знать: современные подходы к проведению статистических расчетов и анализу, статистические пакеты прикладных программ, актуальные научные публикации в области статистического анализа

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
методологии сбора и обработки статистических данных, проводить статистические расчеты с применением соответствующих математических методов и информационных технологий, а также последующую аналитическую работу с полученными данными	применением информационных технологий ИД.ПК-4.3. Анализирует данные на основе методов математической статистики ИД.ПК-4.4. Разрабатывает и совершенствует системы статистических показателей и методик их расчета	З (ПК-4)
		Уметь: производить статистические расчеты с применением соответствующих математических методов и информационных технологий, а также последующую аналитическую работу с полученными данными, применять статистические пакеты прикладных программ У (ПК-4)
		Владеть: навыками применения традиционных и инновационных методов статистического анализа В (ПК-4)

В результате освоения дисциплины обучающийся должен:

**знать:**

- спецификации моделей панельных данных и предпосылки их идентификации;
- методы оценки пространственных моделей и диагностики пространственной зависимости;
- вероятностные основы моделей дискретного выбора и счётных данных и способы интерпретации результатов их оценивания;
- основные понятия моделей длительности состояний.

**уметь:**

- выбирать подходящую модель в зависимости от структуры данных и исследовательской задачи;
- оценивать параметры моделей и проводить проверку статистических гипотез;
- анализировать качество модели и интерпретировать оцененные коэффициенты;
- применять модели для анализа реальных социально-экономических данных.

**владеть:**

- инструментами математической статистики и регрессионного анализа;
- навыками работы с программным обеспечением для эконометрического моделирования;
- методами самостоятельной разработки прикладных моделей прогнозирования.

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина «Панельные, пространственные и дискретные модели» является обязательной дисциплиной части, формируемой участниками образовательных отношений, Блока 1 «Дисциплины (модули)» образовательной программы «Данные, знания, экономика, нарративы». Курс читается в третьем, четвертом модулях, форма промежуточной аттестации – зачет в третьем модуле, зачет с оценкой в четвертом модуле.

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 единицы, 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
Контактная работа обучающихся с преподавателем в соответствии с УП:		42	-	-	21	21	-	-	-	-	-	-
Лекции (Л)		14	-	-	7	7	-	-	-	-	-	-
Семинарские занятия (СЗ)		28	-	-	14	14	-	-	-	-	-	-
Самостоятельная работа (СР)		30	-	-	15	15	-	-	-	-	-	-
Промежуточная аттестация	форма	зачет, зачет с оценкой	-	-	зачет	зачет с оценкой	-	-	-	-	-	-
	час.	-	-	-	-	-	-	-	-	-	-	-
Общая трудоемкость дисциплины (час./з.е.)		72/2	-	-	36/1	36/1	-	-	-	-	-	-

#### 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	Тема 1: Статические панельные модели	Типы панельных данных: сбалансированные и несбалансированные панели. Преимущества и ограничения использования панельных данных. Объединённая модель: спецификация, предпосылки, особенности оценки и интерпретации. Модель с фиксированными эффектами (FE): спецификация, допущения, оценка по отклонениям от индивидуальных средних (within-оценка), интерпретация коэффициентов. Модель со случайными эффектами (RE): предпосылки, обобщённый метод наименьших квадратов (GLS), интерпретация. Сравнение моделей: F-тест на фиксированные эффекты, LM-тест Бройша–Пагана на случайные эффекты, тест Хаусмана для выбора между FE и RE. Модель	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)
		Хаусмана–Тейлора: модель со случайными эффектами, допускающая корреляцию части регрессоров с индивидуальными эффектами. Введение в модели с ограниченной гетерогенностью: модели, допускающие различия между группами.			
2	Тема 2: Динамические модели панельных данных	Спецификация динамической панели: зависимость от лагов зависимой переменной. Проблема смещения оценки при наличии индивидуальных эффектов (смещение Никелла). Мотивация использования инструментальных переменных. Модель Ареллано–Бонда: условия моментной идентификации, выбор инструментов, оценка методом обобщённого метода моментов (GMM). Проверка спецификации: тесты на автокорреляцию остатков первого и второго порядка, тест Хансена на валидность инструментов. Чувствительность к выбору инструментов и их количеству, возможная переидентификация.	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)
3	Тема 3: Пространственная зависимость и её диагностика	Типы пространственных данных и примеры применения. Пространственная зависимость: мотивация учёта в моделях. Пространственная матрица весов: принципы построения, типы (смежность, расстояние, К-ближайших соседей), нормализация. Меры пространственной автокорреляции: глобальный индекс Морана, локальный индекс Морана. Тесты на пространственную зависимость: Тесты множителей Лагранжа на наличие пространственного лага зависимой переменной и пространственной корреляции в ошибке.	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)
4	Тема 4: Пространственные регрессионные модели	Эндогенная пространственная зависимость Модель с пространственным лагом в зависимой переменной (SAR-модель): спецификация, особенности интерпретации, метод оценки. Модель с пространственной корреляцией ошибок (SEM-модель). Модель с пространственными лагами объясняющих переменных (SLX-модель). Интерпретация	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)



№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)
		коэффициентов в пространственных моделях: прямые, косвенные и совокупные эффекты.			
5	Тема 5. Модели дискретного выбора	Общая структура моделей дискретного выбора, метод максимального правдоподобия. Бинарные модели: логит- и пробит-спецификации, интерпретация коэффициентов, предельные эффекты, отношения шансов. Проверка качества модели: псевдо- $R^2$ , точность классификации. Мультиномиальные модели: модель множественного выбора, учёт регрессоров, зависящих от альтернатив; допущение независимости от нерелевантных альтернатив и его ограничения. Модели упорядоченного выбора: индексная модель со скрытой переменной, упорядоченные логит- и пробит-модели, интерпретация порогов и предельных эффектов. Сравнение моделей и выбор подходящей спецификации в зависимости от типа зависимой переменной.	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)
6	Тема 6. Модели счётных данных	Счётные модели: пуассоновская модель — спецификация, свойства, допущение равенства среднего и дисперсии. Отрицательная биномиальная модель: учёт избыточности дисперсии. Модели с избыточным числом нулей: структура с бинарной и счётной компонентой, критерии выбора модели.	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)
7	Тема 7. Основы анализа длительности состояния	Понятие времени до наступления события, цензурирование, функции выживания и риска. Непараметрические методы: оценка Каплана—Мейера (функция выживания), оценка Нельсона—Аалена (кумулятивный риск). Модель пропорционального риска (модель Кокса): спецификация, допущения, интерпретация. Применение моделей длительности в социально-экономических исследованиях.	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)

## 5.2 Структура дисциплины

Таблица 4

### Структура дисциплины

№ п/п	Наименование тем (разделов)	Объем дисциплины, час.					Форма текущего контроля успеваемости*, промежуточно й аттестации
		Всего	Контактная работа обучающихся с преподавателем по типам учебных занятий в соответствии с УП			СР	
			Лек	Лаб	Сем		
Очная форма обучения							
Тема 1	Статические панельные модели.	15	3	-	6	6	ДЗ
Тема 2	Динамические модели панельных данных	10	2	-	4	4	ДЗ
Тема 3	Пространственная зависимость и её диагностика	11	2	-	4	5	ДЗ
Промежуточная аттестация		-	-	-	-	-	-
Всего:		36	7	-	14	15	Зачет
Тема 4	Пространственные регрессионные модели	10	2	-	4	4	ДЗ
Тема 5	Модели дискретного выбора	15	3	-	6	6	ДЗ
Тема 6	Модели счётных данных	5	1	-	2	2	ДЗ
Тема 7	Основы анализа длительности состояния	6	1	-	2	3	ДЗ
Промежуточная аттестация		-	-	-	-	-	Зачет с оценкой
Всего:		72	14	-	14	44	-
Итого:		72	14	-	28	30	-

\*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 соответствующим 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: Статические панельные модели.**

- 1.1. Повторение пройденного на лекциях и на практических занятиях материала – 1 час.
- 1.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 1 час.
- 1.3. Выполнение домашнего задания – 4 часа.
- Итого: 6 часов.

#### **Тема 2: Динамические модели панельных данных**

- 2.1. Повторение пройденного на лекциях и на практических занятиях материала – 1 час.
- 2.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 1 час.
- 2.3. Выполнение домашнего задания – 2 часа.
- Итого: 4 часа.

#### **Тема 3: Пространственная зависимость и её диагностика**

- 3.1. Повторение пройденного на лекциях и на практических занятиях материала – 1 час.
- 3.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 1 час.
- 3.3. Выполнение домашнего задания – 3 часа.
- Итого: 5 часов.

#### **Тема 4. Пространственные регрессионные модели**

- 4.1. Повторение пройденного на лекциях и на практических занятиях материала – 1 час.
- 4.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 1 час.
- 4.3. Выполнение домашнего задания – 2 часа.
- Итого: 4 часа.

#### **Тема 5. Модели дискретного выбора**

- 5.1. Повторение пройденного на лекциях и на практических занятиях материала – 1 час.
- 5.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 1 час.
- 5.3. Выполнение домашнего задания – 4 часа.
- Итого: 6 часов.

#### **Тема 6. Модели счётных данных**

- 6.1. Повторение пройденного на лекциях и на практических занятиях материала – 0,5 часа.
- 6.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 0,5 часа.
- 6.3. Выполнение домашнего задания – 1 час.
- Итого: 2 часа.

#### **Тема 7. Основы анализа длительности состояния**

- 7.1. Повторение пройденного на лекциях и на практических занятиях материала – 0,5 часа.

7.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 0,5 часа.

7.3. Выполнение домашнего задания – 2 часа.

Итого: 3 часа.

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

Тема 1. Статические панельные модели. Понятие панельных данных. Отличие сбалансированных и несбалансированных панелей.

- Спецификация моделей с фиксированными и случайными эффектами: предпосылки, методы оценки, интерпретация коэффициентов.

- Критерии выбора между объединённой моделью, моделями с фиксированными и случайными эффектами.

- Модель Хаусмана–Тейлора: классификация переменных, условия идентификации, внутренняя инструментализация.

Тема 2. Динамические модели панельных данных

- Проблема динамической панели: зависимость от лагов, смещение Никелла, нарушение условий экзогенности.

- Метод Ареллано–Бонда: спецификация модели, обоснование выбора инструментов.

- Диагностика модели: тесты AR(1), AR(2), тест Хансена на валидность инструментов.

Тема 3. Пространственная зависимость и её диагностика (3 вопроса)

- Типы пространственных данных и мотивация учёта пространственной взаимозависимости в моделях.

- Пространственная матрица весов: принципы построения, типы смежности и нормализация.

- Методы диагностики пространственной автокорреляции: глобальный и локальный индекс Морана, LM-тесты.

Тема 4. Пространственные регрессионные модели

- SAR-, SEM- и SLX-модели: спецификации, особенности оценки и различия в интерпретации.

- Эндогенная и экзогенная пространственная зависимость: экономическая и статистическая интерпретация.

- Интерпретация коэффициентов в пространственных моделях: прямые, косвенные и совокупные эффекты.

Тема 5. Модели дискретного выбора

- Бинарные модели: логит и пробит-спецификации, предельные эффекты и интерпретация коэффициентов.

- Мультиномиальные и порядковые модели: спецификации, условия применимости, ограничения.

- Независимость от нерелевантных альтернатив: смысл, последствия нарушения, способы обхода.

Тема 6. Модели счётных данных

- Пуассоновская модель: спецификация, допущения, ограничение равенства среднего и дисперсии.

- Отрицательная биномиальная модель: учёт переизбыточности дисперсии, отличие от Пуассона.

- Модели с избыточным числом нулей: двухкомпонентная структура, предпосылки, интерпретация.

Тема 7. Основы анализа длительности состояния.

- Основные понятия анализа длительности: время до события, правостороннее цензурирование, типы событий.
- Непараметрические методы: оценка функции выживания по Каплану—Мейеру, функции риска по Нельсону—Аалену.
- Модель пропорционального риска (модель Кокса): спецификация, интерпретация, допущение пропорциональности.

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

1. Ратникова, Татьяна Анатольевна. Анализ панельных данных и данных о длительности состояний : учеб. пособ. / Т. А. Ратникова, К. К. Фурманов ; Национальный исследовательский университет "Высшая школа экономики". - Москва : ВШЭ, 2014. - 373 с. - Библиогр. : с. 364 - 373. - ISBN 978-5-7598-1093-3. - Текст : непосредственный.

2. Кэмерон, Колин. Микроэконометрика : методы и их применение : 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. - Москва : Дело, 2015. - (Академический учебник). - ISBN 978-5-7749-0957-5. - Текст : непосредственный. Книга 2. - 2015. - XXVIII, 525-1158 с. - Библиогр. : с. 1082 - 1126. - Пер. изд. : Microeconometrics : Methods and Applications / A. C. Cameron, P. K. Trivedi. - Cambridge ; New York ; Melbourne, 2005. - ISBN 978-5-7749-0956-8. Предм. указ. : с. 1129 - 1151. Имен. указ. : с. 1152 – 1158.

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

Для обеспечения самостоятельной работы магистрантов по дисциплине **«Панельные, пространственные и дискретные модели»** 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 занятию, активное слушание на лекциях, выполнение домашних заданий. Магистрант должен присутствовать на лекциях, отвечать на поставленные вопросы, показывая, что прочитал разбираемую литературу, представлять содержательные реплики по темам обсуждения.

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

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

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Тема 1: Статические панельные модели.	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено
Тема 2: Динамические модели панельных данных	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено
Тема 3: Пространственная зависимость и её диагностика	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено
Тема 4. Пространственные регрессионные модели	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено
Тема 5. Модели дискретного выбора	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено
Тема 6. Модели счётных данных	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено
Тема 7. Основы анализа длительности состояния	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено

При освоении дисциплины домашнее задание оценивается с использованием 100-балльной шкалы с последующим переводом в бинарную систему для получения результатов текущего контроля, фиксирующих ход образовательного процесса, согласно Таблице 6.

Таблица 6

## Критерии оценивания

Формы текущего контроля успеваемости	Описание	Показатели	Количество баллов по 100-балльной шкале	Результаты текущего контроля
Домашнее задание	Выполнение домашнего задания подразумевает письменное изложение ответов на теоретические вопросы и решений поставленных преподавателем задач. Предполагается, что ход решения задач сопровождается подробными комментариями обучающегося. По результатам домашнего задания выставляются баллы.	Обучающийся правильно выполняет задания и отвечает на все поставленные в задании вопросы.	81–100	зачтено
		Обучающийся выполняет предложенные задания и отвечает на все поставленные в задании вопросы, но допускает при этом несущественные ошибки.	61–80	
		Обучающийся предлагает правильную идеологию решения задач.	41–60	
		Обучающийся испытывает затруднения при выполнении предложенных заданий.	0–40	не зачтено

## 7.2 Контрольные задания для текущей аттестации

## Примерный материал домашних заданий

Тема 1: Статические панельные модели.

- Вы — аналитик в исследовательском центре трудовых рынков. Ваша задача — понять, как индивидуальные характеристики (образование, опыт, когнитивные способности) влияют на логарифм заработка  $\ln\_wage$ , используя панельные данные. При этом важно учесть неизмеренные индивидуальные эффекты (например, трудовая мотивация, семейный фон), которые постоянны во времени.

У вас есть файл `wage_panel.csv` с данными:

Имя переменной	Описание переменной
<code>id</code>	идентификатор индивида
<code>time</code>	период наблюдения (индекс волны)
<code>ln_wage</code>	логарифм заработка
<code>educ</code>	годы образования
<code>rexp, rexp2</code>	потенциальный опыт и его квадрат
<code>ability</code>	индекс когнитивной способности
<code>broken_home</code>	индикатор неполной семьи

Задания:

- Оцените объединённую модель.
- Оцените модель с фиксированными эффектами.
- Оцените модель со случайными эффектами.
- Как изменился коэффициент при переменной `educ`?

- e. Проведите необходимые тесты для выбора между объединённой, FE и RE моделями.
- f. Какую модель вы выберите и почему?
- g. На основе полученных оценок дайте рекомендацию стоит ли инвестировать в дополнительное образование с учётом неизмеримых индивидуальных факторов?

2. У вас есть данные: `panel_heterogeneity_data.csv` по нескольким регионам за ряд лет.

Задания:

- a. Постройте модель с фиксированными эффектами по  $i$

$$y_{it} = \alpha_i + \beta x_{it} + u_{it}$$

- b. Для каждого региона  $i$  по отдельности оцените регрессию:

$$y_{it} = \alpha_i + \beta_i x_{it} + u_{it}$$

- Построй гистограмму или диаграмму «ящик с усами» по оценкам  $\beta_i$
- Есть ли визуальные свидетельства гетерогенности?

- c. Разделите регионы на 3 группы по схожести  $\beta_i$ . Постройте модель:

$$y_{it} = \alpha_i + \beta_{g(i)} x_{it} + u_{it}$$

где  $g(i)$  — группа региона.

Сравните оценённые  $\beta_A, \beta_B, \beta_C$  между группами.

- d. Построй модель, позволяющую  $\beta_i$  меняться по годам:

$$y_{it} = \alpha_i + \sum_t \beta_t \times (D\_year_t \times x_{it}) + u_{it}$$

где  $D\_year_t$  — фиктивная переменная для года.

- Построй график траектории  $\beta_t$ .
  - Есть ли тренд?
- e. (дополнительно) Проведите тест на равенство всех  $\beta_i$ .

Тема 2: Динамические модели панельных данных

3. При анализе детерминант заработной платы важно учесть, что текущая зарплата может зависеть от заработной платы в прошлом. Используйте набор данных `wage_panel.csv` с (описание переменных такое же, как в Задании 1).

Задания:



- a. Оцените динамическую модель, в которой текущая заработная плата зависит от собственного лага, а также от индивидуальных факторов. Используйте метод Ареллано–Бонда для устранения смещения и учета индивидуальных эффектов.
- b. Сколько инструментов использовано?
- c. Пройден ли тест на валидность инструментов?
- d. Пройдены ли тесты на автокорреляцию остатков?
- e. Какова оценка коэффициента при лаге  $\ln\_wage_{t-1}$ ? Как это интерпретировать в терминах «инерционности» дохода?
- f. Каков знак и значимость коэффициента при  $educ$ ? Что это говорит о краткосрочном эффекте образования?

### Тема 3: Пространственная зависимость и её диагностика

1. Задания: на сайте Росстата найдите электронное приложение к последнему сборнику «Регионы России. Социально-экономические показатели» за последний доступный год (Главная страница/ Публикации / Каталог публикаций / Статистические издания / Регионы России. Социально-экономические показатели).
  - a. Используйте показатель «Уровень безработицы» (среднегодовые значения) по регионам РФ за предпоследний доступный год.
  - b. Загрузите пространственные геоданные по регионам РФ (субъектам Российской Федерации). Для этого используйте один из следующих источников: в R: shapefile можно получить из пакета `raster` (`raster::regions_map`), либо загрузить с сайта GADM (уровень 1 — субъекты РФ), в Python: используйте библиотеку `geopandas` для загрузки заранее скачанного shapefile с сайта GADM или `geoportal.rgo.ru`. Геоданные должны содержать данные регионов РФ, необходимые для построения пространственной матрицы смежности.
  - c. Постройте пространственную матрицу весов: используйте критерий смежности по границе (регионы считаются соседями, если имеют общую границу). Нормализуйте строки матрицы весов (row-standardization).
  - d. Рассчитайте глобальный индекс Морана для уровня безработицы. Интерпретируйте значение индекса и его статистическую значимость.
  - e. Рассчитайте локальные индексы Морана. Определите регионы, входящие в кластеры высокого и низкого уровня безработицы.
2. Используйте показатели «ВВП на душу населения», «Доля городского населения», «Инвестиции в основной капитал» по регионам РФ за тот же год, что и «Уровень безработицы» доступный год.

#### Задания:

- a. Постройте простую линейную модель объяснения уровня безработицы показателями «ВВП на душу населения», «Доля городского населения», «Инвестиции в основной капитал».
- b. Проведите тесты на пространственную зависимость: на наличие пространственного лага зависимой переменной; на наличие пространственной автокорреляции в остатках.

- с. Определите, какая форма пространственной зависимости (лаг или ошибка) более значима.

#### Тема 4. Пространственные регрессионные модели

1. Используйте показатели «ВВП на душу населения», «Доля городского населения», «Инвестиции в основной капитал» по регионам РФ за тот же год, что и «Уровень безработицы» доступный год.

Задания:

- а. Оцените пространственные модели:
  1. SAR (модель с пространственным лагом зависимой переменной).
  2. SEM (модель с пространственной автокорреляцией ошибок).
  3. SLX (модель с пространственными лагами объясняющих переменных).
- б. Сравните модели:
  - По критерию AIC/BIC.
  - По значимости коэффициентов.
- с. Для SAR-модели:
  - Рассчитайте прямые, косвенные и совокупные эффекты.
  - Объясните, как влияют факторы (например, как рост ВРП в соседнем регионе влияет на безработицу в данном).
- д. Сравните, как учёт пространственной зависимости улучшает объяснение данных.

#### Тема 5. Модели дискретного выбора

1. В файле RussiaINDData.csv содержится подвыборка данных Российского мониторинга экономического положения и здоровья НИУ-ВШЕ.
  - а. Создайте переменные:
    - Индекс массы тела (на основании переменных роста q2507 и веса q2506)
    - Дамми-переменную, равную единице, если человек сейчас курит, и равную нулю, если нет (на основании переменной q3002)
    - Дамми-переменную, равную единице, если человек недавно употреблял алкоголь, и равную, если нет (на основании переменной q3008)
    - На основе переменной q1012 – дамми-переменную, =1 если человек живет с партнером, не важно, в официальном браке или нет, и =0, если человек живет один (даже если он состоит в официальном браке, но проживает с партнером раздельно)
    - Переменную, описывающую наивысший достигнутый уровень образования в трех категориях:
      - «Education: secondary or less»
      - «Education: high school»
      - «Education: college/university or higher»
  - б. Постройте модель бинарного выбора для переменной goodhealth, где, в качестве объясняющих переменных, помимо вновь созданных описанных выше, входят пол, возраст, место проживания (город или село).
  - с. При необходимости удалите из модели совместно незначимые переменные.

- d. Прокомментируйте качество модели.
- e. Проанализируйте, как на вероятность иметь хорошее здоровье влияет курение (в терминах отношения шансов, в терпенах предельных эффектов). Проанализируйте, как на вероятность иметь хорошее здоровье влияет возраст (графически).

#### Тема 6. Модели счётных данных

1. Вы – аналитик в страховой компании и анализируете факторы, влияющие на интенсивность обращения пожилых людей за медицинской помощью. В распоряжении находятся данные о числе визитов к врачу за последний год и различных характеристиках пациентов, включая возраст, наличие страховки, хронические заболевания и т. д.

Ваша задача — оценить и сравнить несколько моделей счётных данных, интерпретировать ключевые эффекты и обосновать выбор наилучшей модели.

#### Описание данных (mus17data.csv)

Переменная	Описание
docvis	<b>Зависимая переменная:</b> число визитов к врачу за год
private	=1, если есть частная медицинская страховка
age	Возраст
age2	Квадрат возраста
educyr	Число лет обучения
actlim	=1, если есть ограничения физической активности
totchr	Число хронических заболеваний

Задания:

- a. Постройте распределение docvis:
  - частотную таблицу,
  - график (гистограмма, желательно с наложением Пуассона).
  - Ответьте: много ли наблюдений с нулевыми визитами? Есть ли избыточность дисперсии?

- b. Оцените пуассоновскую модель:

Интерпретируйте:

- знак и значимость переменных private, actlim, totchr;
- эффект age и age2 (в виде графика или словесно);
- предельные эффекты.

- c. Оцените отрицательную биномиальную модель:

Сравните модели пуассоновскую модель и отрицательную биномиальную модель:

- Диагностирована ли избыточность дисперсии?
- Сравните логарифмы правдоподобия и информационные критерии (AIC/BIC)

- d. Оцените пуассоновскую модель с избыточными нулями, в которой факт наличия страховки объясняет «вероятность участия», а остальные переменные — «интенсивность участия».

Интерпретируйте:

- влияние *private* и на вероятность *необращения* за медицинской помощью,
  - интерпретацию коэффициентов основной части модели.
- е. Оцените отрицательную биномиальную модель с избыточными нулями, в которой факт наличия страховки объясняет «вероятность участия», а остальные переменные – «интенсивность участия».

Сравните оцененные модели, какую модель вы выберите и почему?

- ф. Каков основной вывод о влиянии страхования, хронических заболеваний и ограничений активности на количество визитов?

## Тема 7. Основы анализа длительности состояния

1. Файл *ema1996.csv* содержит следующие данные:

Переменная	Описание
<i>spell</i>	Длительность безработицы в двухнедельных интервалах (1 = 2 недели, 2 = 4 недели и т.д.)
<i>sensor1</i>	=1, если человек вышел из безработицы (найдена работа), =0 если наблюдение было цензурировано
<i>ui</i>	если человек получает пособие по безработице; =0 — не получает

### Задания:

Постройте оценки функции выживания (Каплана—Мейера) для переменной *spell*, стратифицированные по переменной *ui*.

Сравните выживаемость между группами  $ui = 1$  и  $ui = 0$ .

Используйте:

- график функции Каплана—Мейера
- проведите тест логрангов на равенство кривых выживания
- сделайте вывод о статистической значимости различий между группами

Сделайте вывод о влиянии факта получения пособия по безработицы на ее длительность.

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

Форма промежуточной аттестации в третьем модуле – **зачет в форме тестирования**.

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

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

При выставлении оценки по итогам каждого модуля учитываются как результаты тестирования, так и текущего контроля успеваемости. Максимальная **оценка по итогам освоения дисциплины** в каждой промежуточной аттестации составляет 100 баллов.

В 3 модуле 50% оценки составляют результаты выполнения домашних заданий, 50% – результаты зачетного тестирования. В 4 модуле 50 % оценки составляют результаты выполнения домашних заданий, 50% – результаты зачетного тестирования.

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

В рамках тестирования в 3 модуле студентам будет необходимо решить 10 заданий базовой, повышенной и высокой сложности, из которых 4 задания являются заданиями базовой сложности, 4 задания повышенной сложности и 2 задания 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.

В рамках тестирования в 4 модуле студентам будет необходимо решить 20 заданий базовой, повышенной и высокой сложности, из которых 10 заданий являются заданиями базовой сложности, 6 заданий повышенной сложности и 4 задания 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 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 вопроса, даны верные обоснования - 2 балла; ответ имеет незначительные отклонения от эталонного, правильно названы на все запрашиваемые составляющие вопроса, 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 представлены в таблице 7.

Таблица 7

### Критерии оценивания тестирования

Вид промежуточной аттестации	Показатели	Количество баллов
Тест	100-81% правильных ответов	81–100
	80-61% правильных ответов	61–80
	60-41% правильных ответов	41–60
	40% и менее правильных ответов	0–40

На основании набранной суммы баллов выставляется итоговая оценка.

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** в третьем модуле и **зачет с оценкой** в четвертом модуле в соответствии с Таблицей 8

Таблица 8

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

Форма промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)	Оценка по итогам освоения дисциплины (в 100-балльной шкале)	Результаты текущего контроля
<b>зачет</b>	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	81–100	Зачтено
				61–80	
				41–60	
				0–40	Не зачтено
<b>зачет с оценкой</b>	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	81–100	Зачтено, отлично
				61–80	Зачтено, хорошо
				41–60	Зачтено, удовлетворительно
				0–40	Не зачтено, неудовлетворительно

Результаты промежуточного контроля по дисциплине, выраженные в бинарной системе «зачтено», 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 бинарной системе «не зачтено», 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).

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

##### Примеры тестовых заданий для промежуточной аттестации

###### 3 модуль:

###### Задание 1

**Компетенция:** ПК-3

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какова основная причина использования GMM (обобщённого метода моментов) при оценке динамических моделей панельных данных?

1. Для получения точных предсказаний при наличии пропущенных значений
2. Для устранения влияния мультиколлинеарности между регрессорами
3. Для устранения смещения, вызванного корреляцией лагов зависимой переменной с индивидуальными эффектами
4. Для учета гетероскедастичности остатков

**Поле для ответа:**

--	--	--	--

###### Задание 2

**Компетенция:** ПК-3

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой тест используется для проверки наличия автокорреляции второго порядка в остатках модели Ареллано–Бонда?

1. Тест Хаусмана
2. Тест LM Бройша–Пагана
3. Тест AR(2)
4. Тест Вальда

**Поле для ответа:**

--	--	--	--

###### Задание 3

**Компетенция:** ПК-4

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Что из перечисленного является примером пространственной матрицы весов?

1. Таблица сопряжённости между категориальными переменными
2. Матрица корреляций между переменными в модели
3. Матрица смежности, показывающая наличие границ между наблюдениями
4. Матрица расстояний между уровнями значимости в гипотезах

**Поле для ответа:**

--	--	--	--

**Задание 4**

**Компетенция:** ПК-4

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой из следующих индексов используется для оценки степени глобальной пространственной автокорреляции переменной?

1. Индекс Джини
2. Индекс Морана
3. Индекс Херфиндаля
4. Индекс Вальда

**Поле для ответа:**

--	--	--	--

**Задание 5**

**Компетенция:** ПК-3

**Уровень сложности:** повышенный

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

При построении модели с фиксированными эффектами на основе региональной панели вы замечаете, что в результатах, выведенных статистическим пакетом, оценки по некоторым регрессорам отсутствуют. Какие из следующих причин могут объяснять это?

**Варианты ответа:**

1. Эти регрессоры не варьируются во времени
2. Эти регрессоры имеют пропущенные значения
3. Эти регрессоры линейно зависимы от фиксированных эффектов
4. Эти регрессоры выражены в логарифмах
5. Эти регрессоры имеют одинаковое значение во всех регионах в каждый момент времени

**Поле для ответа:**



--	--	--	--	--

### Задание 6

**Компетенция:** ПК-3

**Уровень сложности:** повышенный

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие свойства характерны для модели с фиксированными эффектами (FE)?

**Варианты ответа:**

1. FE-модель устраняет смещение от ненаблюдаемых постоянных во времени индивидуальных эффектов, коррелированных с регрессорами.
2. В FE-модели невозможно оценить влияние переменных, не меняющихся во времени (например, пол, регион).
3. Оценки FE совпадают с pooled OLS, если применить усреднение по группам.
4. FE-модель корректна даже при корреляции индивидуальных эффектов с объясняющими переменными.
5.  $R^2$  в FE-модели всегда выше, чем в модели со случайными эффектами (RE)

**Поле для ответа:**

--	--	--	--	--

### Задание 7

**Компетенция:** ПК-4

**Уровень сложности:** повышенный

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Вы анализируете результаты модели с фиксированными эффектами (FE). Какие из следующих признаков указывают на проблемы в интерпретации или качестве модели?

**Варианты ответа:**

1.  $R^2$ -between (межгрупповая объясняющая способность) близок к нулю.
2. Остатки модели имеют выраженную гетероскедастичность, которую вы не учли при оценивании
3. Коэффициенты при переменных с межгрупповой вариацией статистически незначимы
4. Величина стандартных ошибок существенно возрастает при переходе от т модели со случайными эффектами RE к модели с фиксированными эффектами FE
5. Тест Хаусмана не отвергает гипотезу о корреляции эффектов с регрессорами.

**Поле для ответа:**

--	--	--	--	--

### Задание 8

**Компетенция:** ПК-4

**Уровень сложности:** повышенный

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Вы провели LM-тест Бройша–Пагана и получили значимое значение статистики. Какие выводы и действия на основе этого корректны?

1. Есть основания использовать модель со случайными эффектами (RE).
2. Следует отклонить гипотезу об отсутствии индивидуальных эффектов.
3. Необходимо перейти к модели с фиксированными эффектами (FE).
4. Pooled-модель (без учета панельной структуры) допустима.
5. Результат подтверждает валидность RE-модели.

**Поле для ответа:**

--	--	--	--	--

### Задание 9

**Компетенция:** ПК-3

**Уровень сложности:** высокий

**Инструкция:** внимательно прочитайте описание ситуации и ответьте на все вопросы в свободной форме. В ответе необходимо указать выбранную модель, ожидаемый знак коэффициента и краткое обоснование выбора модели с учётом представленных допущений. Ответ должен быть логически последовательным и продемонстрировать понимание предпосылок применяемых моделей.

**Формулировка задания:**

Вы анализируете панельные данные по 30 регионам России за 6 лет.

- Зависимая переменная: уровень занятости (%);
- Ключевая объясняющая переменная: доля расходов бюджета на программы занятости (% от общего бюджета).

При этом известно, что:

- Расходы на занятость слабо меняются во времени внутри регионов, но сильно различаются между регионами
- Регионы с более высокой долей расходов стабильно демонстрируют более высокую занятость
- При этом эксперты отмечают, что размер расходов часто отражает устойчивые приоритеты региональной политики, которые также влияют на занятость через другие каналы

**Вопросы:**

1. Какую модель следует использовать? (FE, RE, Pooled OLS)
2. Какой ожидается знак коэффициента?
3. Кратко обоснуйте выбор модели

**Поле для ответа:**

1. Модель:
2. Знак:
3. Обоснование:

## Задание 10

**Компетенция:** ПК-4

**Уровень сложности:** высокий

**Инструкция:** Внимательно прочитайте описание ситуации и ответьте на все вопросы в свободной форме. В ответе необходимо указать, какие тесты и процедуры следует провести, что именно они проверяют, и как интерпретировать возможные результаты. Ответ должен быть логически последовательным и продемонстрировать понимание методов анализа данных в динамических панелях.

### Формулировка задания:

Вы оцениваете модель производительности труда на основе панельных данных по фирмам за 10 лет. Зависимая переменная — логарифм производительности. Среди объясняющих переменных — лаги зависимой переменной, инвестиции и численность персонала. Для устранения фиксированных фирменных эффектов и смещения, вызванного включением лагов зависимой переменной, вы используете обобщённый метод моментов (GMM) в первых разностях, то есть спецификацию модели Ареллано–Бонда.

После оценки вы получили следующие результаты:

- $p$ -значение теста на автокорреляцию остатков первого порядка ( $AR(1)$ )  $< 0.001$
- $p$ -значение теста на автокорреляцию второго порядка ( $AR(2)$ )  $= 0.28$
- $p$ -значение теста Хансена на валидность инструментов  $= 0.150$

### Вопросы:

1. Какова цель тестов  $AR(1)$  и  $AR(2)$  в модели Ареллано–Бонда?
2. Как следует интерпретировать полученные значения этих тестов?
3. Что проверяет тест Хансена, и что означает его результат в данной ситуации?
4. Можно ли считать спецификацию модели корректной? Обоснуйте.

### Поле для ответа:

- 1.
- 2.
- 3.
- 4.

## 4 модуль:

### Задание 1

**Компетенция:** ПК-3

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

### Формулировка задания:

Какой из следующих методов позволяет устранить смещение оценок коэффициентов регрессии, возникающее из-за корреляции между ненаблюдаемыми индивидуальными эффектами и объясняющими переменными в панельных данных?

1. Метод главных компонент
2. Метод первых разностей

3. Обычный метод наименьших квадратов
4. Метод стратификации

**Поле для ответа:**

--	--	--	--

## Задание 2

**Компетенция:** ПК-3

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Когда модель со случайными эффектами (RE) предпочтительнее модели с фиксированными эффектами (FE)?

1. Когда индивидуальные эффекты коррелированы с объясняющими переменными
2. Когда количество наблюдений по времени очень мало
3. Когда индивидуальные эффекты некоррелированы с объясняющими переменными
4. Когда объясняющая переменная принимает только бинарные значения

**Поле для ответа:**

--	--	--	--

## Задание 3

**Компетенция:** ПК-3

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какая оценочная процедура чаще всего применяется при оценивании модели с фиксированными эффектами?

1. Обобщённый метод наименьших квадратов
2. Метод наименьших квадратов с удалением индивидуальных средних (within-оценка)
3. Метод двухшаговой оптимизации
4. Метод вероятностного выбора

**Поле для ответа:**

--	--	--	--

## Задание 4

**Компетенция:** ПК-3

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой из следующих подходов чаще всего используется для оценки бинарной модели логит или пробит?

1. Метод наименьших квадратов
2. Метод моментов
3. Метод максимального правдоподобия
4. Метод главных компонент

**Поле для ответа:**

--	--	--	--

**Задание 5**

**Компетенция:** ПК-3

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой из следующих терминов обозначает вероятность «пережить» событие до определённого момента времени в моделях анализа длительности?

1. Функция плотности
2. Функция риска
3. Функция выживания
4. Функция правдоподобия

**Поле для ответа:**

--	--	--	--

**Задание 6**

**Компетенция:** ПК-4

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какие данные считаются «сбалансированной панелью»?

1. Если у каждого индивида данные наблюдаются только один раз
2. Если переменные стандартизированы
3. Если для каждой единицы наблюдения представлены все периоды
4. Если структура ошибок одинакова по всем наблюдениям

**Поле для ответа:**

--	--	--	--

**Задание 7**

**Компетенция:** ПК-4

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какое допущение лежит в основе стандартной пуассоновской модели для счётных данных?

1. Среднее значение и дисперсия случайной величины равны
2. Остатки модели нормально распределены
3. Ошибки модели независимы и гомоскедастичны
4. Наблюдаемые значения принимают только два возможных значения

**Поле для ответа:**

--	--	--	--

### Задание 8

**Компетенция:** ПК-4

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Что измеряет глобальный индекс Морана при анализе пространственных данных?

1. Корреляцию между остатками модели и независимыми переменными
2. Среднее расстояние между наблюдениями на карте
3. Наличие пространственной автокорреляции между значениями переменной в соседних единицах
4. Степень мультиколлинеарности между объясняющими переменными

**Поле для ответа:**

--	--	--	--

### Задание 9

**Компетенция:** ПК-4

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какая из характеристик чаще всего используется для оценки качества бинарной модели логистической регрессии?

1. Коэффициент детерминации  $R^2$
2. Точность классификации (accuracy)
3. Средняя абсолютная ошибка (MAE)
4. Доля пропущенных значений

**Поле для ответа:**

--	--	--	--

### Задание 10

**Компетенция:** ПК-4

**Уровень сложности:** базовый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Что означает термин «правостороннее цензурирование» в моделях длительности?

1. Значение переменной измерено с ошибкой
2. Длительность события зафиксирована точно
3. Время наступления события неизвестно, но превышает наблюдаемый период
4. Событие произошло до начала наблюдения

**Поле для ответа:**

--	--	--	--

### Задание 11

**Компетенция:** ПК-3

**Уровень сложности:** повышенный

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие последствия возникнут, если при наличии корреляции между индивидуальными эффектами и регрессорами использовать объединённую регрессию (pooled OLS) вместо модели с фиксированными эффектами (FE)?

**Варианты ответа:**

1. Оценки коэффициентов окажутся смещёнными и несостоятельными
2. Стандартные ошибки оценок будут вычислены неверно
3. Коэффициенты можно интерпретировать как средние эффекты для всех объектов
4. Значение  $R^2$  будет искусственно завышено
5. Причинно-следственные выводы сохранят достоверность

**Поле для ответа:**

--	--	--	--	--

### Задание 12

**Компетенция:** ПК-3

**Уровень сложности:** повышенный

**Инструкция:**

Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Вы строите регрессионную модель на основе панельных данных для анализа влияния региональных инвестиций на ВРП. Какие из следующих утверждений свидетельствуют о

необходимости использования модели с фиксированными эффектами (FE), а не объединённой регрессии (Pooled OLS)?

1. В данных присутствуют ненаблюдаемые индивидуальные характеристики регионов, не включённые в модель, но влияющие на зависимую переменную.
2. Индивидуальные эффекты предполагаются некоррелированными с регрессорами.
3. Оценки объединённой модели (Pooled OLS) существенно отличаются от within-оценок (FE-модели).
4. Результат F-теста на значимость фиксированных эффектов оказывается статистически значимым.
5. Модель построена по сбалансированной панели.

**Поле для ответа:**

--	--	--	--	--

### Задание 13

**Компетенция:** ПК-3

**Уровень сложности:** повышенный

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Для корректного применения SAR-модели (модели с пространственным лагом зависимой переменной) необходимо:

1. Убедиться, что пространственные связи имеют строгую иерархию (например, только столица влияет на регионы).
2. Построить пространственную матрицу весов.
3. Проверить наличие пространственной автокорреляции остатков в обычной регрессии.
4. Использовать метод максимального правдоподобия или GMM для оценки.
5. Убедиться, что все регионы имеют одинаковое число соседей.

**Поле для ответа:**

--	--	--	--	--

### Задание 14

**Компетенция:** ПК-4

**Уровень сложности:** повышенный

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

В каком случае для прогнозирования регионального ВРП по панельным данным предпочтительна модель со случайными эффектами (RE)?

**Варианты ответа:**

1. RE-модель всегда предпочтительнее FE-модели при малом количестве временных периодов (T) и большом числе объектов (N), независимо от корреляции индивидуальных эффектов с регрессорами



2. Все объясняющие переменные варьируются только во времени
3. Индивидуальные эффекты не коррелируют с регрессорами
4. Тест Хаусмана даёт незначимый результат
5. FE-модель невозможно оценить по техническим причинам

**Поле для ответа:**

--	--	--	--	--

### Задание 15

**Компетенция:** ПК-4

**Уровень сложности:** повышенный

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие действия приводят к некорректной интерпретации результатов модели с фиксированными эффектами?

1. Интерпретация коэффициента при переменной, не меняющейся во времени (например, постоянный признак региона)
2. Игнорирование коррелированных с ошибкой индивидуальных эффектов
3. Использование нестандартизированной переменной в логарифмической форме
4. Отсутствие учета панельной структуры при оценке стандартных ошибок
5. Нормирование всех переменных перед оценкой FE-модели

**Поле для ответа:**

--	--	--	--	--

### Задание 16

**Компетенция:** ПК-4

**Уровень сложности:** повышенный

**Формулировка задания:**

Вы используете мультиномиальную логит-модель (MNL), где зависимая переменная — тип используемой финансовой услуги (вклад, кредит, страхование). Что необходимо учесть при интерпретации результатов?

1. Коэффициенты задают эффект по сравнению с базовой категорией
2. Коэффициенты не имеют прямой интерпретации в вероятностях
3. Предполагается независимость от нерелевантных альтернатив (IIA)
4. Модель допускает корреляцию между альтернативами
5. Предельные эффекты одинаковы для всех наблюдений

**Поле для ответа:**

--	--	--	--	--

### Задание 17

**Компетенция:** ПК-3

**Уровень сложности:** высокий

**Инструкция:** Прочитайте описание задачи и выберите наиболее подходящую модель. Обоснуйте выбор и кратко опишите интерпретацию коэффициентов.

**Формулировка задания:**

Вы исследуете влияние инвестиций в культуру на доход на душу населения по 60 регионам. Известно:

- доход в регионе зависит от доходов в соседних регионах
- инвестиции в культуру существенно варьируются по регионам
- данные включают переменные: инвестиции, урбанизация, численность населения

**Вопросы:**

1. Какую модель пространственной регрессии целесообразно использовать (SAR, SEM, SLX)?
2. Отражает ли коэффициент при переменной «инвестиции в культуру» общий эффект влияния инвестиций на доход данного региона в этой модели?

**Поле для ответа:**

- 1.
- 2.

### Задание 18

**Компетенция:** ПК-3

**Уровень сложности:** высокий

**Инструкция:** Внимательно прочитайте описание ситуации и ответьте в свободной форме.

**Формулировка задания:**

Исследуется выбор студентами направления магистратуры. Варианты выбора: экономика, финансы, маркетинг, менеджмент. Данные включают индивидуальные характеристики (баллы ЕГЭ, доход семьи, пол) и сведения об альтернативах (количество бюджетных мест, средняя зарплата выпускников).

Вы решили использовать мультиномиальную логит-модель (MNL).

**Вопросы:**

1. Почему мультиномиальная логит-модель подходит для этой задачи?
2. Что означает предпосылка независимости от нерелевантных альтернатив (IIA)? Почему она может быть проблемной в данной ситуации?

**Поле для ответа:**

- 1.
- 2.

### Задание 19

**Компетенция:** ПК-4

**Уровень сложности:** высокий

**Инструкция:** Внимательно прочитайте описание ситуации и ответьте в свободной форме.

В ответе необходимо обосновать выбор модели, описать её структуру и предложить процедуры для проверки качества и пригодности модели с точки зрения обработки статистических данных.

**Формулировка задания:**

Вы анализируете данные об использовании поликлинических услуг: переменная — количество визитов пациента за год. В выборке большая доля наблюдений равна нулю (пациенты не обращались вовсе), много значений 1–2, но также встречаются значения 10 и более визитов. Оценка пуассоновской регрессии показывает, что дисперсия существенно превышает среднее значение переменной, а также фиксируется избыток нулей по сравнению с предсказаниями модели.

**Вопросы:**

1. Почему пуассоновская модель может быть некорректна в данной ситуации?
2. Какую альтернативную спецификацию модели целесообразно использовать? Опишите её структуру и эконометрическую мотивацию.

**Поле для ответа:**

- 1.
- 2.

**Задание 20****Компетенция:** ПК-4**Уровень сложности:** высокий

**Инструкция:** Прочитайте описание ситуации. Ответьте на вопросы по интерпретации результатов и диагностике модели.

**Формулировка задания:**

Вы оцениваете стандартную линейную регрессию методом наименьших квадратов. По итогам оценки:

- все коэффициенты при регрессорах значимы и интерпретируемы;
- глобальный индекс Морана, рассчитанный для остатков модели, оказался статистически значимым ( $p < 0.01$ );
- LM-тест на пространственную автокорреляцию ошибок также показал значимость ( $p < 0.05$ ).

**Вопросы:**

1. Какая модель пространственной регрессии предпочтительна и почему?
2. Как проверить, устранена ли пространственная зависимость после оценки модели?

**Поле для ответа:**

- 1.
- 2.

**7.5 Средства оценки индикаторов достижения компетенций**

Таблица 9

**Средства оценки индикаторов достижения компетенций**

Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Средства оценки (в соот. с Таблицами 5, 7)
ПК-3	ИД.ПК-3.1. ИД.ПК-3.2.	домашнее задание (ДЗ), тестирование
ПК-4	ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	домашнее задание (ДЗ), тестирование

Таблица 10

## Описание средств оценки индикаторов достижения компетенций

Средства оценки (в соот. с Таблицами 5, 7)	Рекомендованный план выполнения работы
Домашнее задание	<p>Магистрант в ходе подготовки и выполнения домашнего задания по темам дисциплины, показывает способность совершать следующий набор профессиональных действий, получивший развитие в рамках данной дисциплины:</p> <p>— Составляет прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом на основании самостоятельно разработанных моделей с использованием математических, статистических и инструментальных методов, а также современных интеллектуальных технологий</p> <p>— Применяет математические методы и информационные технологии, проводит статистические расчеты, работает с различными источниками статистической информации, ведет аналитическую работу с данными, разрабатывает и совершенствует методологии сбора и обработки статистических данных</p>
Тестирование	<p>Магистрант в ходе подготовки и выполнения тестирования по темам дисциплины, показывает способность совершать следующий набор профессиональных действий, получивший развитие в рамках данной дисциплины:</p> <p>— Составляет прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом на основании самостоятельно разработанных моделей с использованием математических, статистических и инструментальных методов, а также современных интеллектуальных технологий</p> <p>— Применяет математические методы и информационные технологии, проводит статистические расчеты, работает с различными источниками статистической информации, ведет аналитическую работу с данными, разрабатывает и совершенствует методологии сбора и обработки статистических данных</p>

## 8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

### 8.1. Основная литература

1. Кэмерон, Колин. Микроэконометрика : методы и их применение : 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. - Москва : Дело, 2015. - (Академический учебник). - ISBN 978-5-7749-0957-5. - Текст : непосредственный. Книга 2. - 2015. - XXVIII, 525-1158 с. - Библиогр. : с. 1082 - 1126. - Пер. изд. : Microeconometrics : Methods and Applications / A. C. Cameron, P. K. Trivedi. - Cambridge ; New York ; Melbourne, 2005. - ISBN 978-5-7749-0956-8. Предм. указ. : с. 1129 - 1151. Имен. указ. : с. 1152 – 1158.

2. Ратникова, Татьяна Анатольевна. Анализ панельных данных и данных о длительности состояний : учеб. пособ. / Т. А. Ратникова, К. К. Фурманов ; Национальный исследовательский университет "Высшая школа экономики". - Москва : ВШЭ, 2014. - 373 с. - Библиогр. : с. 364 - 373. - ISBN 978-5-7598-1093-3. - Текст : непосредственный.

### 8.2. Дополнительная литература

1. Сяо Чэн. Анализ панельных данных : учебник / Сяо Чэн ; пер. с англ. под ред. В. Сидоренко ; Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации. – Москва : Дело, 2022. – 624 с. : ил. – (Академический учебник). – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=694960>. – Библиогр.: с. 549-587. – ISBN 978-5-85006-381-8. – Текст : электронный..

## **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
14. Python
15. Stata

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

#### **Информационно-справочные системы**

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

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

1. Национальная электронная библиотека НЭБ: <http://www.rusneb.ru>
2. Президентская библиотека: <http://www.prlib.ru>
3. Российская государственная библиотека: <http://www.rsl.ru/>
4. 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:

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

#### **Электронные библиотечные системы:**

1. **Znanium.com** – Электронная библиотечная система (ЭБС) – <http://znanium.com/>;
2. 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, лицензионные электронные ресурсы библиотеки Университета, официальный сайт Университета (Европейский университет в Санкт-Петербурге [<https://euspr.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).

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ**  
**«Панельные, пространственные и дискретные модели»**



## 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 занятию, активное слушание на лекциях, выполнение домашних заданий. Магистрант должен присутствовать на лекциях, отвечать на поставленные вопросы, показывая, что прочитал разбираемую литературу, представлять содержательные реплики по темам обсуждения.

Текущий контроль проводится в форме оценивания выполнения 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 тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Тема 1: Статические панельные модели.	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено
Тема 2: Динамические модели панельных данных	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено
Тема 3: Пространственная зависимость и её диагностика	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено
Тема 4. Пространственные регрессионные модели	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено
Тема 5. Модели дискретного выбора	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Тема 6. Модели счётных данных	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено
Тема 7. Основы анализа длительности состояния	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	Домашние задания	зачтено / не зачтено

При освоении дисциплины домашнее задание оценивается с использованием 100-балльной шкалы с последующим переводом в бинарную систему для получения результатов текущего контроля, фиксирующих ход образовательного процесса, согласно Таблице 2.

Таблица 2

#### Критерии оценивания

Формы текущего контроля успеваемости	Описание	Показатели	Количество баллов по 100-балльной шкале	Результаты текущего контроля
Домашнее задание	Выполнение домашнего задания подразумевает письменное изложение ответов на теоретические вопросы и решений поставленных преподавателем задач. Предполагается, что ход решения задач сопровождается подробными комментариями обучающегося. По результатам домашнего задания выставаются баллы.	Обучающийся правильно выполняет задания и отвечает на все поставленные в задании вопросы.	81–100	зачтено
		Обучающийся выполняет предложенные задания и отвечает на все поставленные в задании вопросы, но допускает при этом несущественные ошибки.	61–80	
		Обучающийся предлагает правильную идеологию решения задач.	41–60	
		Обучающийся испытывает затруднения при выполнении предложенных заданий.	0–40	не зачтено

## 2 Контрольные задания для текущей аттестации

### Материал домашних заданий

Тема 1: Статические панельные модели.

1. Вы — аналитик в исследовательском центре трудовых рынков. Ваша задача — понять, как индивидуальные характеристики (образование, опыт, когнитивные способности) влияют на логарифм заработка  $\ln\_wage$ ,

используя панельные данные. При этом важно учесть неизмеренные индивидуальные эффекты (например, трудовая мотивация, семейный фон), которые постоянны во времени.

У вас есть файл `wage_panel.csv` с данными:

Имя переменной	Описание переменной
<code>id</code>	идентификатор индивида
<code>time</code>	период наблюдения (индекс волны)
<code>ln_wage</code>	логарифм заработка
<code>educ</code>	годы образования
<code>rexp, rexp2</code>	потенциальный опыт и его квадрат
<code>ability</code>	индекс когнитивной способности
<code>broken home</code>	индикатор неполной семьи

Задания:

- Оцените объединённую модель.
- Оцените модель с фиксированными эффектами.
- Оцените модель со случайными эффектами.
- Как изменился коэффициент при переменной `educ`?
- Проведите необходимые тесты для выбора между объединённой, FE и RE моделями.
- Какую модель вы выберете и почему?
- На основе полученных оценок дайте рекомендацию стоит ли инвестировать в дополнительное образование с учётом неизмеримых индивидуальных факторов?

2. У вас есть данные: `panel_heterogeneity_data.csv` по нескольким регионам за ряд лет.

Задания:

- Постройте модель с фиксированными эффектами по  $i$

$$y_{it} = \alpha_i + \beta x_{it} + u_{it}$$

- Для каждого региона  $i$  по отдельности оцените регрессию:

$$y_{it} = \alpha_i + \beta_i x_{it} + u_{it}$$

- Построй гистограмму или диаграмму «ящик с усами» по оценкам  $\beta_i$
- Есть ли визуальные свидетельства гетерогенности?

- Разделите регионы на 3 группы по схожести  $\beta_i$ . Постройте модель:

$$y_{it} = \alpha_i + \beta_{g(i)} x_{it} + u_{it}$$

где  $g(i)$  — группа региона.

Сравните оценённые  $\beta_A, \beta_B, \beta_C$  между группами.

d. Построй модель, позволяющую  $\beta_i$  меняться по годам:

$$y_{it} = \alpha_i + \sum_t \beta_t \times (D\_year_t \times x_{it}) + u_{it}$$

где  $D\_year_t$  — фиктивная переменная для года.

- Построй график траектории  $\beta_t$ .
- Есть ли тренд?

e. (дополнительно) Проведите тест на равенство всех  $\beta_i$ .

## Тема 2: Динамические модели панельных данных

1. При анализе детерминант заработной платы важно учесть, что текущая зарплата может зависеть от зарплаты в прошлом. Используйте набор данных `wage_panel.csv` с (описание переменных такое же, как в Задании 1).

Задания:

- a. Оцените динамическую модель, в которой текущая заработная плата зависит от собственного лага, а также от индивидуальных факторов. Используйте метод Ареллано–Бонда для устранения смещения и учета индивидуальных эффектов.
- b. Сколько инструментов использовано?
- c. Пройден ли тест на валидность инструментов?
- d. Пройдены ли тесты на автокорреляцию остатков?
- e. Какова оценка коэффициента при лаге  $\ln\_wage_{t-1}$ ? Как это интерпретировать в терминах «инерционности» дохода?
- f. Каков знак и значимость коэффициента при  $educ$ ? Что это говорит о краткосрочном эффекте образования?

## Тема 3: Пространственная зависимость и её диагностика

1. Задания: на сайте Росстата найдите электронное приложение к последнему сборнику «Регионы России. Социально-экономические показатели» за последний доступный год (Главная страница/ Публикации / Каталог публикаций / Статистические издания / Регионы России. Социально-экономические показатели).
- a. Используйте показатель «Уровень безработицы» (среднегодовые значения) по регионам РФ за предпоследний доступный год.
- b. Загрузите пространственные геоданные по регионам РФ (субъектам Российской Федерации). Для этого используйте один из следующих источников: в R: `shapefile` можно получить из пакета `rusmap` (`rusmap::regions_map`), либо загрузить с сайта GADM (уровень 1 — субъекты РФ), в Python: используйте библиотеку `geopandas` для загрузки заранее скачанного `shapefile` с сайта GADM или `geoportal.rgo.ru`. Геоданные должны содержать данные регионов РФ, необходимые для построения пространственной матрицы смежности.

- с. Постройте пространственную матрицу весов: используйте критерий смежности по границе (регионы считаются соседями, если имеют общую границу). Нормализуйте строки матрицы весов (row-standardization).
  - d. Рассчитайте глобальный индекс Морана для уровня безработицы. Интерпретируйте значение индекса и его статистическую значимость.
  - e. Рассчитайте локальные индексы Морана. Определите регионы, входящие в кластеры высокого и низкого уровня безработицы.
2. Используйте показатели «ВВП на душу населения», «Доля городского населения», «Инвестиции в основной капитал» по регионам РФ за тот же год, что и «Уровень безработицы» доступный год.

Задания:

- a. Постройте простую линейную модель объяснения уровня безработицы показателями «ВВП на душу населения», «Доля городского населения», «Инвестиции в основной капитал».
- b. Проведите тесты на пространственную зависимость: на наличие пространственного лага зависимой переменной; на наличие пространственной автокорреляции в остатках.
- с. Определите, какая форма пространственной зависимости (лаг или ошибка) более значима.

#### Тема 4. Пространственные регрессионные модели

1. Используйте показатели «ВВП на душу населения», «Доля городского населения», «Инвестиции в основной капитал» по регионам РФ за тот же год, что и «Уровень безработицы» доступный год.

Задания:

- a. Оцените пространственные модели:
  1. SAR (модель с пространственным лагом зависимой переменной).
  2. SEM (модель с пространственной автокорреляцией ошибок).
  3. SLX (модель с пространственными лагами объясняющих переменных).
- b. Сравните модели:
  - По критерию AIC/BIC.
  - По значимости коэффициентов.
- с. Для SAR-модели:
  - Рассчитайте прямые, косвенные и совокупные эффекты.
  - Объясните, как влияют факторы (например, как рост ВРП в соседнем регионе влияет на безработицу в данном).
- d. Сравните, как учёт пространственной зависимости улучшает объяснение данных.

#### Тема 5. Модели дискретного выбора

1. В файле RussiaINDData.csv содержится подвыборка данных Российского мониторинга экономического положения и здоровья НИУ-ВШЕ.

а. Создайте переменные:

- Индекс массы тела (на основании переменных роста q2507 и веса q2506)
  - Дамми-переменную, равную единице, если человек сейчас курит, и равную нулю, если нет (на основании переменной q3002)
  - Дамми-переменную, равную единице, если человек недавно употреблял алкоголь, и равную, если нет (на основании переменной q3008)
  - На основе переменной q1012 – дамми-переменную, =1 если человек живет с партнером, не важно, в официальном браке или нет, и =0, если человек живет один (даже если он состоит в официальном браке, но проживает с партнером отдельно)
  - Переменную, описывающую наивысший достигнутый уровень образования в трех категориях:
    - «Education: secondary or less»
    - «Education: high school»
    - «Education: college/university or higher»
- б. Постройте модель бинарного выбора для переменной goodhealth, где, в качестве объясняющих переменных, помимо вновь созданных описанных выше, входят пол, возраст, место проживания (город или село).
- с. При необходимости удалите из модели совместно незначимые переменные.
- д. Прокомментируйте качество модели.
- е. Проанализируйте, как на вероятность иметь хорошее здоровье влияет курение (в терминах отношения шансов, в терпенах предельных эффектов). Проанализируйте, как на вероятность иметь хорошее здоровье влияет возраст (графически).

## Тема 6. Модели счётных данных

1. Вы – аналитик в страховой компании и анализируете факторы, влияющие на интенсивность обращения пожилых людей за медицинской помощью. В распоряжении находятся данные о числе визитов к врачу за последний год и различных характеристиках пациентов, включая возраст, наличие страховки, хронические заболевания и т. д.

Ваша задача — оценить и сравнить несколько моделей счётных данных, интерпретировать ключевые эффекты и обосновать выбор наилучшей модели.

### Описание данных (mus17data.csv)

Переменная	Описание
docvis	<b>Зависимая переменная:</b> число визитов к врачу за год
private	=1, если есть частная медицинская страховка
age	Возраст
age2	Квадрат возраста
educyr	Число лет обучения
actlim	=1, если есть ограничения физической активности
totchr	Число хронических заболеваний

Задания:

- a. Постройте распределение `docvis`:
- частотную таблицу,
  - график (гистограмма, желательно с наложением Пуассона).
  - Ответьте: много ли наблюдений с нулевыми визитами? Есть ли избыточность дисперсии?
- b. Оцените пуассоновскую модель:
- Интерпретируйте:
- знак и значимость переменных `private`, `actlim`, `totchr`;
  - эффект `age` и `age2` (в виде графика или словесно);
  - предельные эффекты.
- c. Оцените отрицательную биномиальную модель:
- Сравните модели пуассоновскую модель и отрицательную биномиальную модель:
- Диагностирована ли избыточность дисперсии?
  - Сравните логарифмы правдоподобия и информационные критерии (AIC/BIC)
- d. Оцените пуассоновскую модель с избыточными нулями, в которой факт наличия страховки объясняет «вероятность участия», а остальные переменные – «интенсивность участия».
- Интерпретируйте:
- влияние `private` и на вероятность *необращения* за медицинской помощью,
  - интерпретацию коэффициентов основной части модели.
- e. Оцените отрицательную биномиальную модель с избыточными нулями, в которой факт наличия страховки объясняет «вероятность участия», а остальные переменные – «интенсивность участия».
- Сравните оцененные модели, какую модель вы выберите и почему?
- f. Каков основной вывод о влиянии страхования, хронических заболеваний и ограничений активности на количество визитов?

## Тема 7. Основы анализа длительности состояния

1. Файл `ema1996.csv` содержит следующие данные:

Переменная	Описание
<code>spell</code>	Длительность безработицы в двухнедельных интервалах (1 = 2 недели, 2 = 4 недели и т.д.)
<code>sensor1</code>	=1, если человек вышел из безработицы (найдена работа), =0 если наблюдение было цензурировано
<code>ui</code>	если человек получает пособие по безработице; =0 — не получает

### Задания:

Постройте оценки функции выживания (Каплана—Мейера) для переменной `spell`, стратифицированные по переменной `ui`.

Сравните выживаемость между группами  $ui = 1$  и  $ui = 0$ .

Используйте:

- график функции Каплана—Мейера

- проведите тест логрангов на равенство кривых выживания
- сделайте вывод о статистической значимости различий между группами

Сделайте вывод о влиянии факта получения пособия по безработицы на ее длительность.

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

Форма промежуточной аттестации в третьем модуле – **зачет в форме тестирования.**

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

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

При выставлении оценки по итогам каждого модуля учитываются как результаты тестирования, так и текущего контроля успеваемости. Максимальная **оценка по итогам освоения дисциплины** в каждой промежуточной аттестации составляет 100 баллов.

В 3 модуле 50% оценки составляют результаты выполнения домашних заданий, 50% – результаты зачетного тестирования. В 4 модуле 50 % оценки составляют результаты выполнения домашних заданий, 50% – результаты зачетного тестирования.

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

В рамках тестирования в 3 модуле студентам будет необходимо решить 10 заданий базовой, повышенной и высокой сложности, из которых 4 задания являются заданиями базовой сложности, 4 задания повышенной сложности и 2 задания 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.

В рамках тестирования в 4 модуле студентам будет необходимо решить 20 заданий базовой, повышенной и высокой сложности, из которых 10 заданий являются заданиями базовой сложности, 6 заданий повышенной сложности и 4 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 вопроса, даны верные обоснования - 2 балла; ответ имеет незначительные отклонения от эталонного, правильно названы на все запрашиваемые составляющие вопроса, 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 представлены в таблице 3.

Таблица 3

#### **Критерии оценивания тестирования**

Вид промежуточной аттестации	Показатели	Количество баллов
Тест	100-81% правильных ответов	81–100
	80-61% правильных ответов	61–80
	60-41% правильных ответов	41–60
	40% и менее правильных ответов	0–40

На основании набранной суммы баллов выставляется итоговая оценка.

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** в третьем модуле и **зачет с оценкой** в четвертом модуле 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 соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)	Оценка по итогам освоения дисциплины (в 100-балльной шкале)	Результаты текущего контроля
зачет	ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4)	81–100	Зачтено
				61–80	
				41–60	
				0–40	Не зачтено

Форма промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)	Оценка по итогам освоения дисциплины (в 100-балльной шкале)	Результаты текущего контроля
		ИД.ПК-4.3. ИД.ПК-4.4.	У (ПК-4) В (ПК-4)		
зачет оценкой	с ПК-3 ПК-4	ИД.ПК-3.1. ИД.ПК-3.2. ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	З (ПК-3) У (ПК-3) В (ПК-3) З (ПК-4) У (ПК-4) В (ПК-4)	81–100	Зачтено, отлично
				61–80	Зачтено, хорошо
				41–60	Зачтено, удовлетворительно
				0–40	Не зачтено, неудовлетворительно

Результаты промежуточного контроля по дисциплине, выраженные в бинарной системе «зачтено», 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 бинарной системе «не зачтено», 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).

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

##### Перечень тестовых заданий для промежуточной аттестации

**ПК-3 Способен составлять прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом на основании самостоятельно разработанных моделей с использованием математических, статистических и инструментальных методов, а также современных интеллектуальных технологий**

#### Задания базового уровня сложности

##### Задание 1

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой из следующих признаков отличает модель с фиксированными эффектами от объединённой регрессии?

1. Использование инструментальных переменных
2. Наличие случайных ошибок с кластерной структурой
3. Учет индивидуальных констант для каждой единицы панели
4. Использование переменных, изменяющихся только между группами

**Поле для ответа:**

--	--	--	--

## Задание 2

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой из следующих методов позволяет устранить смещение, возникающее из-за корреляции индивидуальных эффектов и объясняющих переменных в панельной модели?

1. Метод главных компонент
2. Метод первых разностей
3. Обычный метод наименьших квадратов
4. Метод стратификации

**Поле для ответа:**

--	--	--	--

## Задание 3

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой тест применяется для проверки значимости фиксированных эффектов по сравнению с объединённой регрессией

1. Тест Дарбина—Уотсона
2. Тест Хаусмана
3. F-тест на значимость групповых эффектов
4. Тест Вальда

**Поле для ответа:**

--	--	--	--

## Задание 4

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какая процедура оценки чаще всего используется для модели с фиксированными эффектами?

1. Обобщённый метод наименьших квадратов
2. Метод наименьших квадратов с удалением индивидуальных средних (within-оценка)
3. Метод двухшаговой оптимизации
4. Метод вероятностного выбора

**Поле для ответа:**

--	--	--	--

**Задание 5**

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Когда модель со случайными эффектами (RE) предпочтительнее модели с фиксированными эффектами (FE)?

1. Когда индивидуальные эффекты коррелированы с объясняющими переменными
2. Когда количество наблюдений по времени очень мало
3. Когда индивидуальные эффекты не коррелированы с объясняющими переменными
4. Когда объясняющая переменная принимает только бинарные значения

**Поле для ответа:**

--	--	--	--

**Задание 6**

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой подход используется в модели Хаусмана–Тейлора для устранения корреляции между регрессорами и индивидуальными эффектами?

1. Использование инструментов
2. Исключение индивидуальных эффектов путём первых разностей в модели фиксированных эффектов
3. Усреднение панельных наблюдений по пятилетним группам
4. Модификация матрицы весов в пространственной модели

**Поле для ответа:**

--	--	--	--

**Задание 7**

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой из следующих методов используется для оценки параметров пробит-модели дискретного выбора?

1. Метод наименьших квадратов
2. Метод моментов
3. Метод максимального правдоподобия
4. Метод главных компонент

**Поле для ответа:**

--	--	--	--

**Задание 8**

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой из следующих показателей применим для оценки качества бинарной модели дискретного выбора?

1. R-квадрат (коэффициент детерминации)
2. Псевдо- $R^2$
3. AIC, но не псевдо- $R^2$
4. Только отношение правдоподобия

**Поле для ответа:**

--	--	--	--

**Задание 9**

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какое ключевое допущение необходимо для корректного применения мультиномиальной логит-модели?

1. Гомоскедастичность ошибок
2. Независимость от нерелевантных альтернатив (ПА)
3. Нормальность ошибок
4. Равенство вероятностей всех альтернатив

Поле для ответа:

--	--	--	--

### Задание 10

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какая функция в анализе длительности показывает вероятность «дожить» до определённого момента времени без наступления события?

1. Функция плотности
2. Функция риска
3. Функция выживания
4. Функция правдоподобия

Поле для ответа:

--	--	--	--

### Задание 11

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой тип модели лучше всего подходит для прогнозирования вероятности участия домохозяйства в программе субсидирования?

1. Линейная регрессия
2. Мультиномиальная логит-модель
3. Бинарная пробит-модель
4. Модель продолжительности

Поле для ответа:

--	--	--	--

### Задание 12

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

При моделировании количества визитов пациента в поликлинику за месяц 40% наблюдений равны нулю. Какую модель целесообразнее использовать?

1. Линейная регрессия
2. Модель Пуассона
3. Модель Пуассона с избыточными нулями
4. Бинарная логит-модель

**Поле для ответа:**

--	--	--	--

### Задание 13

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

В модели упорядоченного выбора коэффициент при переменной «образование» равен +0.4. Что это означает?

1. Вероятность каждой категории увеличивается на 0.4
2. Увеличение образования увеличивает вероятность перехода в более высокую категорию
3. Образование влияет только на пороговые значения
4. Коэффициент показывает, на сколько лет продлевается образование

**Поле для ответа:**

--	--	--	--

### Задание 14

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какая модель лучше всего подходит для оценки влияния уровня безработицы на длительность поиска работы?

1. Модель логит
2. Модель линейной регрессии
3. Модель Каплана—Мейера
4. Модель Кокса

**Поле для ответа:**

--	--	--	--

### Задание 15

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

При прогнозировании вероятности того, что клиент останется в банке, исследователь получает предельный эффект переменной «доход» равный 0.015. Как его интерпретировать?

1. Увеличение дохода на единицу повышает вероятность ухода клиента на 1.5%
2. Увеличение дохода на единицу повышает вероятность остаться на 1.5 процентных пункта
3. Увеличение дохода в два раза повышает шансы остаться в банке
4. Эффект не подлежит интерпретации без отношения шансов

**Поле для ответа:**

--	--	--	--

### Задание 16

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Что лежит в основе вычисления предельных эффектов в бинарных моделях?

1. Производная вероятности по регрессору
2. Сравнение логарифмов правдоподобия
3. Деление на стандартную ошибку
4. Стандартизованное среднее

**Поле для ответа:**

--	--	--	--

### Задание 17

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**



С каким типом зависимой переменной используется мультиномиальная логит-модель?

1. Категориальная с более чем двумя значениями
2. Число событий за период времени
3. Бинарная
4. Длительность состояния до события

**Поле для ответа:**

--	--	--	--

### Задание 18

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Как интерпретируются коэффициенты в логит-модели?

1. Как средние значения зависимой переменной
2. Как отношение шансов (odds ratio)
3. Как отклонения от среднего
4. Как доли вероятности

**Поле для ответа:**

--	--	--	--

### Задание 19

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Что обозначает пороговое значение в упорядоченной пробит-модели?

1. Значение коэффициента перед регрессором
2. Верхний предел случайной ошибки
3. Раздел между латентными значениями, соответствующими разным категориям
4. Уровень значимости теста

**Поле для ответа:**

--	--	--	--

### Задание 20

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какую модель следует использовать, если зависимая переменная принимает три и более неупорядоченных значения?

1. Бинарная логит-модель
2. Мультиномиальная логит-модель
3. Упорядоченная пробит-модель
4. Линейная регрессия

**Поле для ответа:**

--	--	--	--

**Задание 21**

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой из критериев позволяет сравнивать модели дискретного выбора с разным числом параметров, учитывая баланс качества подгонки и сложности модели?

1. t-статистика
2. Р-значение
3. AIC (информационный критерий Акаике)
4. Коэффициент детерминации

**Поле для ответа:**

--	--	--	--

**Задание 22**

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какова основная причина использования GMM (обобщённого метода моментов) при оценке динамических моделей панельных данных?

1. Для получения точных предсказаний при наличии пропущенных значений
2. Для устранения влияния мультиколлинеарности между регрессорами
3. Для устранения смещения, вызванного корреляцией лагов зависимой переменной с индивидуальными эффектами
4. Для учета гетероскедастичности остатков

**Поле для ответа:**

--	--	--	--

### Задание 23

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Почему оценка с фиксированными эффектами (FE) в динамической модели панельных данных может быть смещённой?

1. Из-за гетероскедастичности в ошибках
2. Из-за корреляции индивидуальных эффектов с лагом зависимой переменной
3. Потому что не включены регрессоры
4. Из-за мультиколлинеарности между объясняющими переменными

**Поле для ответа:**

--	--	--	--

### Задания повышенного уровня сложности

#### Задание 1.

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие последствия возникнут, если при наличии корреляции между индивидуальными эффектами и регрессорами использовать объединённую регрессию (pooled OLS) вместо модели с фиксированными эффектами (FE)?

**Варианты ответа:**

1. Оценки коэффициентов окажутся смещёнными и несостоятельными
2. Стандартные ошибки оценок будут вычислены неверно
3. Коэффициенты можно интерпретировать как средние эффекты для всех объектов
4. Значение  $R^2$  будет искусственно завышено
5. Причинно-следственные выводы сохранят достоверность

**Поле для ответа:**

--	--	--	--	--

#### Задание 2

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие свойства характерны для модели с фиксированными эффектами (FE)?

**Варианты ответа:**

1. FE-модель устраняет смещение от неизмеряемых (ненаблюдаемых) постоянных во времени индивидуальных эффектов, коррелированных с регрессорами
2. В FE-модели невозможно оценить влияние переменных, не меняющихся во времени (например, пол, регион)
3. Оценки FE совпадают с pooled OLS, если применить усреднение по группам
4. FE-модель корректна даже при корреляции индивидуальных эффектов с объясняющими переменными
5.  $R^2$  в FE-модели всегда выше, чем в модели со случайными эффектами (RE)

**Поле для ответа:**

--	--	--	--	--

**Задание 3**

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

В каком случае для прогнозирования регионального ВРП по панельным данным предпочтительна модель со случайными эффектами (RE), по сравнению с моделью с фиксированными эффектами (FE)?

**Варианты ответа:**

1. RE-модель всегда предпочтительнее FE-модели при малом количестве временных периодов (T) и большом числе объектов (N), независимо от корреляции индивидуальных эффектов с регрессорами
2. Все объясняющие переменные варьируются только во времени
3. Индивидуальные эффекты не коррелируют с регрессорами
4. Тест Хаусмана не выявляет статистически значимой разницы между оценками моделей FE и RE
5. FE-модель невозможно оценить по техническим причинам

**Поле для ответа:**

--	--	--	--	--

**Задание 4**

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

При построении модели с фиксированными эффектами на основе региональной панели вы замечаете, что в результатах, выведенных статистическим пакетом, оценки по некоторым регрессорам отсутствуют. Какие из следующих причин могут объяснять это?

**Варианты ответа:**

1. Эти регрессоры не варьируются во времени
2. . Эти регрессоры исключаются из оценки, если хотя бы в одном наблюдении есть пропуск
3. Эти регрессоры линейно зависимы от фиксированных эффектов
4. Эти регрессоры выражены в логарифмах
5. Эти регрессоры имеют одинаковое значение во всех регионах в каждый момент времени

**Поле для ответа:**

--	--	--	--	--

**Задание 5**

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Вы строите статическую панельную модель для оценки влияния демографических и экономических характеристик на уровень бедности в субъектах Российской Федерации. Какие из следующих утверждений о моделях с фиксированными и случайными эффектами являются корректными в этом контексте?

**Варианты ответа:**

1. Только модель с фиксированными эффектами остаётся состоятельной при наличии корреляции между неизмеряемыми региональными характеристиками и объясняющими переменными
2. При наличии корреляции между индивидуальными эффектами и регрессорами предпочтительна модель со случайными эффектами
3. Модель случайных эффектов эффективнее, чем модель фиксированных эффектов, если предпосылка некоррелированности выполнена
4. Регрессоры без временной вариации можно оценить только в модели с фиксированными эффектами
5. Модель с фиксированными эффектами не позволяет оценить влияние переменных, которые не меняются во времени внутри наблюдаемой единицы

**Поле для ответа:**

--	--	--	--	--

**Задание 6**

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:**

Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Вы строите регрессионную модель на основе панельных данных для анализа влияния региональных инвестиций на ВРП. Какие из следующих утверждений свидетельствуют о необходимости использования модели с фиксированными эффектами (FE), а не объединённой регрессии (Pooled OLS)?

1. В данных присутствуют ненаблюдаемые индивидуальные характеристики регионов, не включённые в модель, но влияющие на зависимую переменную.
2. Индивидуальные эффекты предполагаются некоррелированными с регрессорами.
3. Оценки объединённой модели (Pooled OLS) существенно отличаются от within-оценок (FE-модели).
4. Результат F-теста на значимость фиксированных эффектов оказывается статистически значимым.
5. Модель построена по сбалансированной панели.

**Поле для ответа:**

--	--	--	--	--

**Задание 7**

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие из следующих утверждений верны для модели Ареллано–Бонда?

1. Требуется строгая экзогенность всех регрессоров
2. Требуется нормальное распределение ошибок
3. На первом этапе применяется преобразование первых разностей для устранения индивидуальных фиксированных эффектов
4. Предполагает отсутствие автокорреляции остатков второго порядка
5. Оценивается с помощью обобщённого метода моментов (GMM)

**Поле для ответа:**

--	--	--	--	--

**Задание 8**

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Дайте краткую интерпретацию результатов оценивания модели Ареллано–Бонда.

**Формулировка задания:**

Вы применили модель Ареллано–Бонда к панели фирм. Коэффициент при лаге прибыли равен 0.6 и значим, а коэффициенты при текущих инвестициях и размере фирмы — положительны и значимы.

Что можно сказать о динамике прибыли и роли других факторов?

Поле для ответа: \_\_\_\_\_

### Задание 9

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие меры помогут избежать свержидентификации (переидентификации) в модели Ареллано–Бонда?

1. Ограничение числа лагов, используемых как инструменты
2. Использование системы GMM
3. Агрегация инструментов
4. Увеличение временного горизонта панели
5. Введение пространственной матрицы весов

Поле для ответа:

--	--	--	--	--

### Задание 10

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:**

Внимательно прочитайте описание и дайте краткий обоснованный ответ, интерпретируя коэффициенты пространственной модели.

**Формулировка задания:**

В модели с пространственным лагом зависимой переменной (SAR) оценено:

$$\log(\text{Цена}) = 0.6 \cdot W \cdot \log(\text{Цена}) + 0.4 \cdot \log(\text{Доход}) + \dots$$

Интерпретируйте значение коэффициента при пространственном лаге зависимой переменной. Какое влияние он отражает?

Поле для ответа: \_\_\_\_\_

### Задание 11

**Компетенция:** ПК-3

**Тип задания:** открытый

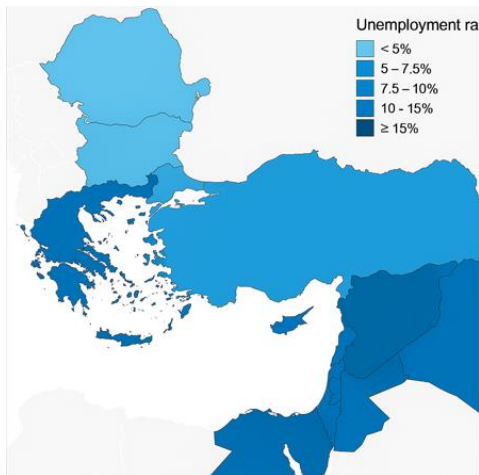
**Инструкция:** Изучите карту и результат пространственного анализа. Ответьте кратко и обоснованно.

**Формулировка задания:**

На карте показан уровень безработицы в регионах Восточной Европы. По результатам расчёта глобального индекса Морана для этой переменной получено значение  $I=0.37$ ,  $p\text{-value} = 0.004$ .

На карте видно, что регионы с высокой безработицей (темнее) формируют кластер в юго-восточной части.

Сделайте интерпретацию пространственного паттерна и статистического вывода.



**Поле для ответа:** \_\_\_\_\_

### Задание 12

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Прочитайте результат оценки и сделайте интерпретацию.

**Формулировка задания:**

Оценена SAR-модель, в которой логарифм цены жилья используется в качестве зависимой переменной, а среди объясняющих переменных — логарифм дохода населения.

Результаты:

- Коэффициент при пространственном лаге зависимой переменной:  $\rho=0.4$ ,  $p = 0.003$
- Коэффициент при  $\log(\text{доход})$ :  $\beta = 0.3$ ,  $p = 0.002$
- Прямой эффект дохода: 0.25
- Косвенный эффект дохода: 0.15
- Совокупный эффект дохода: 0.40

Как интерпретировать прямой, косвенный и совокупный эффект дохода?

**Поле для ответа:** \_\_\_\_\_

### Задание 13

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:**

Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Вы оцениваете влияние уровня образования на уровень занятости в муниципалитетах с учётом пространственной зависимости. Вы построили SAR-модель:

$$y = \rho W y + X \beta + \varepsilon$$

Где  $y$  — уровень занятости,  $X$  — уровень образования и другие контрольные переменные,  $W$  — пространственная матрица весов.

Какие из следующих утверждений являются верными?

1. Коэффициент  $\beta$  напрямую интерпретируется как влияние образования без поправки на пространственную зависимость
2. Коэффициент  $\rho$  отражает силу прямого влияния образования на занятость



3. Прямой эффект отражает влияние образования в данном регионе на занятость в том же регионе
4. Косвенный эффект показывает, как образование в других регионах влияет на занятость в данном регионе
5. Совокупный эффект равен сумме прямого и косвенного эффектов

**Поле для ответа:**

--	--	--	--	--

#### Задание 14

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие из следующих подходов корректны при интерпретации коэффициентов в упорядоченной логит-модели?

**Варианты ответа:**

1. Анализ направления влияния объясняющих переменных на латентную переменную
2. Интерпретация коэффициентов через изменение логарифма кумулятивного отношения шансов для перехода в более высокую категорию
3. Использование коэффициентов как оценок вероятностей
4. Расчёт предельных эффектов по категориям исходной переменной
5. Интерпретация коэффициентов как изменения условных средних значений

**Поле для ответа:**

--	--	--	--	--

#### Задание 15

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Вы построили логит-модель для прогнозирования выбора потребителей (покупать/не покупать) и хотите оценить её качество.

Выберите все методы, которые можно использовать для оценки качества модели.

**Варианты ответа:**

1. Доля правильных классификаций
2. Логарифм функции правдоподобия
3. График остатков по прогнозам
4. Псевдо- $R^2$  (например,  $R^2$  МакФаддена)
5. Проверка нормальности ошибок

**Поле для ответа:**

--	--	--	--	--

### Задание 16

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Внимательно прочитайте описание модели и сделайте вывод, какие из утверждений корректны. Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Изучается влияние курения матери (переменная *smoking* = 1 если курит, 0 если нет) на риск смерти ребёнка с помощью логит-регрессии. Полученный коэффициент при *smoking*: 0.5 ( $p\text{-value} < 0.05$ ).

Какие из следующих утверждений корректны?

**Варианты ответа:**

1. У детей курящих матерей шансы смерти выше, чем у детей некурящих.
2. Курение матери увеличивает вероятность смерти ребёнка на 50%.
3. По модели, у курящих матерей шансы смерти ребёнка примерно в 1.65 раза выше ( $e^{0.5} \approx 1.65$ ) раза.
4. Логит-модель предполагает линейную зависимость между переменной *smoking* и вероятностью смерти.
5. Коэффициент 0.5 можно интерпретировать как изменение вероятности смерти при переходе от 0 к 1.
6. Чтобы рассчитать точное изменение вероятности, нужно вычислить предельные эффекты.

**Поле для ответа:**

--	--	--	--	--	--

### Задание 17

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Внимательно прочитайте описание модели и сделайте вывод, какие из утверждений корректны. Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Изучается влияние уровня образования (переменная *education*) на вероятность карьерного роста (категории: низкий/средний/высокий) с помощью упорядоченной пробит-модели. Получен положительный и статистически значимый коэффициент при *education*.

Какие из следующих интерпретаций верны?

**Варианты ответа:**

1. Люди с более высоким образованием чаще переходят в высшие карьерные категории

2. Влияние образования проявляется только на крайние категории зависимой переменной
3. При росте уровня образования значения латентной переменной  $Y^*$  «потенциал карьерного роста». увеличивается.
4. Знак коэффициента можно интерпретировать аналогично линейной регрессии.
5. Пороговые значения модели зависят от переменной education.

**Поле для ответа:**

--	--	--	--	--

### Задание 18

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные утверждения и отметьте их в соответствующих клетках поля для ответа..

**Формулировка задания:**

Исследование влияния дохода на вероятность владения смартфоном при помощи логит-модели показало, что:

При доходе 50 тыс. руб. вероятность владения  $\approx 60\%$

При доходе 60 тыс. руб. вероятность владения  $\approx 68\%$

Доход включён в модель как количественная переменная (в тыс. руб.)

Оценённый коэффициент при доходе равен 0.7 и статистически значим.

Выберите все верные интерпретации результатов оценивания.

**Варианты ответа:**

1. Увеличение дохода связано с ростом вероятности владения смартфоном.
2. Вероятность владения растёт линейно с доходом: каждые +10 тыс. руб. дают фиксированное увеличение на 8 п.п.
3. Отношение шансов владения при доходах 51 и 50 тыс. составляет примерно 2 ( $\exp(0.7) \approx 2.01$ ).
4. Вероятность владения смартфоном растёт одинаково при любом уровне дохода
5. Влияние дохода на вероятность владения зависит от уровня дохода

**Поле для ответа:**

--	--	--	--	--

### Задание 19

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите все корректные утверждения и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

В модели оценки числа запросов на государственные субсидии (переменная applications) используется отрицательная биномиальная модель. Коэффициент при переменной unemployment\_rate равен 0.3.

Выберите все верные интерпретации результатов оценивания.

1. Повышение уровня безработицы связано с ростом ожидаемого числа заявок на субсидии.
2. При росте безработицы на 1 п.п. число заявок увеличивается на 0.3.
3. Логарифм ожидаемого числа заявок увеличивается на 0.3 при росте безработицы на 1 п.п.
4. IRR составляет примерно 1.35 ( $e^{0.3} \approx 1.35$ ), что означает рост ожидаемого количества обращений примерно на 35% при росте безработицы на 1 п.п.
5. Коэффициент 0.3 можно напрямую интерпретировать как среднее абсолютное изменение количества заявок.

**Поле для ответа:**

--	--	--	--	--

## Задание 20

**Компетенция:** ПК-3

**Тип задания:** закрытый

**Инструкция:** Выберите все корректные утверждения и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

В отрицательной биномиальной модели, объясняющей число налоговых проверок компаний (переменная *audits*), коэффициент при переменной *firm\_size* (размер компании) равен  $-0.4$ .

Какие из следующих утверждений верны?

1. Более крупные компании в среднем подвергаются меньшему числу проверок
2. IRR составляет примерно 0.67 ( $e^{-0.4} \approx 0.67$ ),  $IRR = e^{-0.4} \approx 0.67$  что указывает на снижение ожидаемого количества проверок примерно на 33% при увеличении *firm\_size* на одну единицу
3. Логарифм ожидаемого числа проверок увеличивается при росте размера компании.
4. Коэффициент  $-0.4$  указывает на снижение количества проверок на 0.4 при увеличении *firm\_size* на одну единицу
5. IRR — это отношение шансов между двумя группами, как в логит-модели.

**Поле для ответа:**

--	--	--	--	--

## Задания высокого уровня сложности

### Задание 1

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** внимательно прочитайте описание ситуации и ответьте на все вопросы в свободной форме. В ответе необходимо указать выбранную модель, ожидаемый знак коэффициента и краткое обоснование выбора модели с учётом представленных допущений.

Ответ должен быть логически последовательным и продемонстрировать понимание предпосылок применяемых моделей.

**Формулировка задания:**

Вы анализируете панельные данные по 30 регионам России за 6 лет.

- Зависимая переменная: уровень занятости (%);
- Ключевая объясняющая переменная: доля расходов бюджета на программы занятости (% от общего бюджета).

При этом известно, что:

- Регионы с более высокой долей расходов стабильно демонстрируют более высокую занятость
- При этом эксперты отмечают, что размер расходов часто отражает устойчивые, неизменные во времени, приоритеты региональной политики, которые также влияют на занятость через другие каналы

**Вопросы:**

1. Какую модель следует использовать – с фиксированными или со случайными эффектами?
2. Какой ожидается знак коэффициента?
3. Кратко обоснуйте выбор модели

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.
- 3.

**Задание 2**

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** На основе описания задачи укажите, какая модель является наиболее подходящей.

**Формулировка задания:**

Вы анализируете влияние плотности населения на уровень заработной платы в регионах на основе панельных данных за 15 лет. Предполагается, что в регионах есть неизменяемые особенности, влияющие на заработную плату (например, культура или история), которые не меняются во времени. Какую модель следует использовать для получения состоятельных оценок и почему?

**Поле для ответа:** \_\_\_\_\_

**Задание 3**

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Дайте краткий ответ в свободной форме.

**Формулировка задания:**

В оценённой вами в статистическом пакете модели фиксированных эффектов переменная «пол» (мужчина = 1, женщина = 0) исключается из спецификации автоматически. Почему это происходит и как интерпретировать этот результат?

**Поле для ответа:** \_\_\_\_\_

#### **Задание 4**

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Дайте краткий ответ, объясняющий эконометрическую логику выбора модели.

**Формулировка задания:**

Вы анализируете данные по 200 компаниям за 8 лет. Цель — оценить влияние объёма экспорта на прибыль. Предполагается, что существуют постоянные отличия между компаниями (например, качество менеджмента), которые не наблюдаются напрямую, но могут влиять и на экспорт, и на прибыль. Какую модель следует использовать и почему?

**Поле для ответа:** \_\_\_\_\_

#### **Задание 5**

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Объясните, к каким последствиям может привести игнорирование индивидуальных эффектов.

**Формулировка задания:**

Исследователь использует объединенную модель (pooled OLS) на панельных данных, игнорируя возможное наличие индивидуальных эффектов  $\alpha_i$ . Объясните, к каким последствиям для оценок коэффициентов и выводов исследования может привести игнорирование индивидуальных эффектов при использовании объединённой модели.

**Поле для ответа:** \_\_\_\_\_

#### **Задание 6**

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Проанализируйте проблему и обоснуйте выбор модели.

**Формулировка задания:**

Вы оцениваете влияние уровня образования в прошлом году на текущий уровень дохода в региональной панельной выборке (30 регионов, 10 лет). Образование считается медленно меняющимся показателем. Вы включили лаг дохода и лаг образования как объясняющие переменные.

**Вопросы:**

1. Почему стандартная фиксированная модель (FE) может дать несостоятельные оценки в данной ситуации?
2. Как решает эту проблему модель Ареллано–Бонда?
3. Какие инструменты она использует?

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.
- 3.

### Задание 7

**Компетенция:** ПК-3

**Тип задания:** открытый

**Формулировка задания:**

Исследование изучает влияние прошлой занятости на текущий доход, используя панельные данные домохозяйств за 8 лет. Известно, что:

- Текущий доход сильно коррелирует с доходом в предыдущих периодах.
- Уровень образования (неизменный во времени) также оказывает влияние на доход.

**Вопросы:**

1. Объясните, почему в данной ситуации нельзя использовать обычную регрессию с фиксированными эффектами.
2. Какую модель следует применить вместо нее? Обоснуйте выбор.
3. Укажите возможный источник смещения в оценках и объясните его природу.

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.
- 3.

### Задание 8

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задачи и выберите наиболее подходящую модель. Обоснуйте выбор и кратко опишите интерпретацию коэффициентов.

**Формулировка задания:**

Вы исследуете влияние инвестиций в культуру на доход на душу населения по 60 регионам. Известно:

- доход в регионе зависит от доходов в соседних регионах
- инвестиции в культуру существенно варьируются по регионам
- данные включают переменные: инвестиции, урбанизация, численность населения

**Вопросы:**

1. Какую модель пространственной регрессии целесообразно использовать (SAR, SEM, SLX)?
2. Отражает ли коэффициент при переменной «инвестиции в культуру» общий эффект влияния инвестиций на доход данного региона в этой модели?

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.

### Задание 9

**Тип задания:** открытый

**Инструкция:** Прочитайте описание и сформулируйте ответ в свободной форме.

**Формулировка задания:**

Исследователь оценивает влияние качества дорог (по шкале 1–5) на уровень занятости в муниципалитетах. Используется SAR-модель. Получены следующие оценки:

- Коэффициент при переменной «качество дорог»: 0.08

- Прямой эффект: 0.07
- Косвенный эффект: 0.06
- Совокупный эффект: 0.13

Объясните, как следует интерпретировать прямой и косвенный эффекты. Почему они отличаются от исходного коэффициента? Какие дополнительные расчёты необходимы для интерпретации результатов пространственной модели?

**Поле для ответа:** \_\_\_\_\_

## Задание 10

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Прочитайте описание прикладной задачи. Ответьте в свободной форме, обоснуйте выбор модели и диагностических процедур.

**Формулировка задания:**

Вы анализируете различия в уровне безработицы между 70 регионами России по данным за один год. В начальной линейной регрессии использованы следующие объясняющие переменные:

- доля занятых в промышленности,
- уровень образования,
- средняя заработная плата.

Результаты модели:

- все коэффициенты значимы и интерпретируемы,
- остатки демонстрируют значимую пространственную автокорреляцию (глобальный индекс Морана значим на уровне 1%),
- LM-тест выявляет пространственную зависимость в ошибке,

**Вопросы:**

Какую модель пространственной регрессии целесообразно использовать в данной прикладной задаче? Обоснуйте выбор.

**Поле для ответа:** \_\_\_\_\_

## Задание 11

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Внимательно прочитайте описание ситуации и ответьте в свободной форме.

**Формулировка задания:**

Исследуется выбор студентами направления магистратуры. Варианты выбора: экономика, финансы, маркетинг, менеджмент. Данные включают индивидуальные характеристики (баллы ЕГЭ, доход семьи, пол) и сведения об альтернативах (количество бюджетных мест, средняя зарплата выпускников).

Вы решили использовать мультиномиальную логит-модель (MNL).

**Вопросы:**

1. Почему мультиномиальная логит-модель подходит для этой задачи?
2. Что означает предпосылка независимости от нерелевантных альтернатив (ПА)? Почему она может быть проблемной в данной ситуации?

**Поле для ответа:** \_\_\_\_\_



- 1.
- 2.

### Задание 12

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задачи. Сформулируйте ответ в свободной форме, с акцентом на выбор модели и интерпретацию коэффициентов.

**Формулировка задания:**

Опрос оценивает удовлетворённость граждан качеством городской инфраструктуры по шкале от 1 до 5. Исследователь использует упорядоченную логит-модель, включающую переменные: возраст, доход, тип занятости, размер города.

**Вопросы:**

- Почему упорядоченная логит-модель подходит для этого случая?
- Как интерпретируется коэффициент при переменной доход?
- Что можно сделать, если в крайних категориях слишком мало наблюдений?

**Поле для ответа:** \_\_\_\_\_

### Задание 13

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задачи. Обоснуйте выбор модели и прокомментируйте интерпретацию коэффициентов.

**Формулировка задания:**

Вы анализируете участие студентов в программе обмена. Оценены две модели: логит и пробит.

В логит-модели:

- коэффициент при переменной «уверенность в знании английского» равен **0.8**,
- отношение шансов (odds ratio)  $\approx$  **2.2**.

В пробит-модели:

- соответствующий коэффициент равен **0.45**.

Исследователь утверждает, что интерпретация результатов одинакова для логит- и пробит-моделей.

**Вопросы:**

1. Является ли утверждение исследователя корректным?
2. Как интерпретировать коэффициент 0.8 в логит-модели?
3. Почему логит-модель предпочтительна, если интересуется интерпретация в терминах отношения шансов?
4. Как соотносятся коэффициенты из логит- и пробит-моделей?

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.
- 3.
- 4.

#### Задание 14

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задачи. Обоснуйте выбор модели и предложите процедуру проверки корректности.

**Формулировка задания:**

Изучается количество дней, когда домохозяйства пользовались телемедицинскими консультациями за год. Многие из них не пользовались вовсе, небольшая часть — 20 и более раз. Мы заметили, что остаточная дисперсия сильно превышает среднее.

Вопросы:

- Будет ли пуассоновская модель давать некорректные результаты?
- Какую альтернативную модель следует применить?

**Поле для ответа:** \_\_\_\_\_

#### Задание 15

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задачи. Обоснуйте, в чём различие двух подходов к расчёту предельного эффекта, и прокомментируйте их интерпретацию.

**Формулировка задания:**

При оценке участия в накопительной пенсионной системе исследователь построил модель, которая включает переменную «высшее образование» (бинарная: 1 = есть высшее образование, 0 = нет). Эта переменная значима и положительна в обеих моделях.

По результатам исследователь рассчитал:

- Средний предельный эффект (Average Marginal Effect, AME) в логит-модели = 0.12
- Предельный эффект в средней точке (Marginal Effect at the Mean, MEM) = 0.07

**Вопросы:**

1. Как интерпретируются значения 0.12 и 0.07 в контексте этой модели?
2. Почему они отличаются?
3. В каких ситуациях предпочтительнее использовать средний предельный эффект?

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.
- 3

#### Задание 16

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задачи. Рассчитайте предельный эффект бинарной переменной и интерпретируйте результат.

**Формулировка задания:**

Исследуется влияние наличия смартфона на вероятность участия в онлайн-обучении.

Пробит-модель включает переменную *smartphone* (1 — есть, 0 — нет) и другие контрольные переменные.

Исследователь сообщает:

- При средних значениях остальных переменных, вероятность участия составляет:
  - 0.65 для владельцев смартфона
  - 0.42 для тех, у кого смартфона нет

### Вопросы:

Каков предельный эффект переменной *smartphone* в этой модели? Как интерпретировать полученную величину? Почему важно указывать, в какой точке рассчитан эффект?

Поле для ответа: \_\_\_\_\_

## Задание 17

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задачи. Рассчитайте предельный эффект бинарной переменной, интерпретируйте результат.

### Формулировка задания:

Рассматривается влияние проживания в городе (*urban* = 1, если городской житель) на вероятность обращения за профилактической медицинской помощью. Модель — логит, в неё входят *urban*, *age*, *education*, *income*. Исследователь сообщает:

- При средних значениях переменных *age*, *education*, *income*:
  - вероятность обращения составляет 0.58 для городских жителей
  - 0.41 для сельских жителей

### Вопросы:

1. Каков предельный эффект переменной *urban*?
2. Какова интерпретация этой разницы в терминах вероятностей?
3. Почему предельный эффект может быть другим в другой точке пространства признаков?

Поле для ответа: \_\_\_\_\_

- 1.
- 2.
- 3

## Задание 18

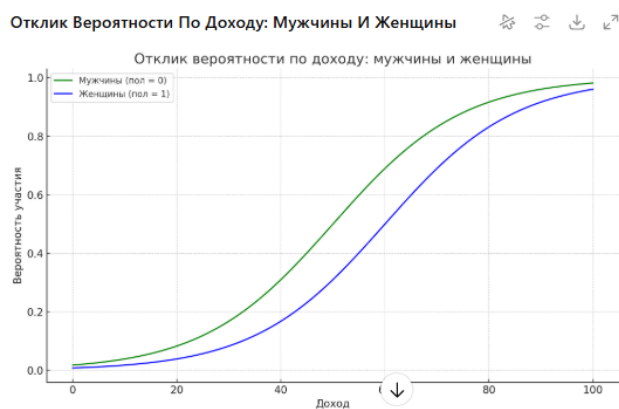
**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задания и ответьте на его вопросы.

### Формулировка задания:

Изучите график ниже, построенный по результатам пробит-модели вероятности участия в программе (ось *Y* — вероятность, ось *X* — доход). Модель включает бинарную переменную «женщина», значение которой 1 для женщин и 0 для мужчин.



На графике показаны две кривые отклика вероятности:

- нижняя — женщины (пол = 1),
- верхняя — мужчины (пол = 0).

**Вопросы:**

1. Какой знак имеет коэффициент при переменной «женщина»?
2. Как в этом графике можно интерпретировать предельный эффект переменной «женщина»?
3. Почему предельный эффект переменной «доход» меняется вдоль оси X?

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.
- 3.

## Задание 19

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задания и ответьте на его вопросы.

**Формулировка задания:**

Вы оцениваете упорядоченную логит-модель, где зависимая переменная — удовлетворённость работой (1 = неудовлетворён, 2 = нейтрально, 3 = удовлетворён).

Коэффициент при переменной «гибкий график» равен 0.9, значим на уровне 5%

**Вопросы:**

Что означает это значение? Как интерпретировать его в терминах латентной переменной?

**Поле для ответа:** \_\_\_\_\_

## Задание 20

**Компетенция:** ПК-3

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задания и ответьте на его вопросы.

**Формулировка задания:**

Вы оцениваете логит-модель вероятности смены работы. В модель включены следующие переменные:

- female: бинарная переменная (1 = женщина, 0 = мужчина),
- married: бинарная переменная (1 = состоит в браке, 0 = не состоит),

- взаимодействие  $\text{female} \times \text{married}$ .

Результаты оценки:

- Коэффициент при  $\text{female} = -0.30$ ,
- Коэффициент при  $\text{married} = -0.20$ ,
- Коэффициент при  $\text{female} \times \text{married} = -0.50$ .

**Вопросы:**

1. Как интерпретируется коэффициент при переменной  $\text{female} \times \text{married}$  в этой модели?
2. Во сколько раз отличаются отношения шансов смены работы у замужних женщин по сравнению с неженатыми мужчинами?

Поле для ответа: \_\_\_\_\_

- 1.
- 2.

**ПК-4 Способен работать с различными источниками статистической информации, разрабатывать и совершенствовать методологии сбора и обработки статистических данных, проводить статистические расчеты с применением соответствующих математических методов и информационных технологий, а также последующую аналитическую работу с полученными данными**

**Задания базового уровня сложности**

**Задание 1**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой тест используется для проверки необходимости учета случайных эффектов в модели панельных данных (по сравнению с объединенной регрессией)?

1. F-тест
2. LM-тест Бройша—Пагана
3. Тест Хаусмана
4. Тест Уайта

**Поле для ответа:**

--	--	--	--

**Задание 2**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Вы провели тест Хаусмана для сравнения моделей с фиксированными (FE) и случайными эффектами (RE). Тест показал, что нулевая гипотеза отвергается на уровне значимости 5% ( $p\text{-value} < 0.05$ ). Что это означает?

1. Модель RE предпочтительнее, так как она более эффективна
2. Не удастся выбрать между FE и RE
3. Предпочтительнее использовать FE, так как допущения RE нарушены
4. Модель RE не требует коррекции на индивидуальные эффекты

**Поле для ответа:**

--	--	--	--

### Задание 3

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какие данные считаются «сбалансированной панелью»?

1. Если у каждого индивида данные наблюдаются только один раз
2. Если переменные стандартизированы
3. Если для каждой единицы наблюдения представлены все периоды
4. Если структура ошибок одинакова по всем наблюдениям

**Поле для ответа:**

--	--	--	--

### Задание 4

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой из признаков говорит о том, что модель с фиксированными эффектами может быть предпочтительнее модели со случайными эффектами?

1. У регрессоров высокая мультиколлинеарность
2. Тест Харке-Бера показал, что остатки модели не нормальны на уровне значимости 5% ( $p\text{-value} < 0.05$ )
3. Тест Хаусмана показал, что нулевая гипотеза отвергается на уровне значимости 5% ( $p\text{-value} < 0.05$ )
4. Структура панели несбалансирована

**Поле для ответа:**

--	--	--	--

### Задание 5

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

При анализе панельных данных мы провели LM-тест Бреуша–Пагана (на случайные эффекты), результаты теста показали, что нулевая гипотеза не отвергается ( $p\text{-value} > 0.05$ ). Какой вывод из этого следует?

1. Следует выбрать модель со случайными эффектами (RE)
2. Можно использовать объединённую регрессию (pooled OLS)
3. Необходимо применить модель с инструментальными переменными
4. Требуется перейти к динамической панельной модели

**Поле для ответа:**

--	--	--	--

**Задание 6**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какую цель преследует F-тест при сравнении модели с фиксированными эффектами и объединённой модели?

1. Проверка наличия автокорреляции
2. Проверка валидности инструментов
3. Проверка значимости индивидуальных (групповых) фиксированных эффектов
4. Проверка гетероскедастичности

**Поле для ответа:**

--	--	--	--

**Задание 7**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

В чём заключается ключевая проблема оценки динамической панели с фиксированными эффектами методом обычных наименьших квадратов?

**Варианты ответов:**

1. Оценка невозможна при большом числе временных периодов
2. Возникает смещение Никелла из-за корреляции лагов зависимой переменной с индивидуальными эффектами
3. Ошибки становятся гетероскедастичными
4. Пропущенные переменные приводят к мультиколлинеарности

**Поле для ответа:**

--	--	--	--

### Задание 8

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой из следующих тестов используется для проверки валидности инструментов в модели Ареллано–Бонда?

**Варианты ответов:**

1. Тест Хаусмана
2. Тест Дарбина–Уотсона
3. Тест Хансена
4. LM-тест Бройша–Пагана

**Поле для ответа:**

--	--	--	--

### Задание 9

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой тест используется для проверки наличия автокорреляции второго порядка в остатках модели Ареллано–Бонда?

1. Тест Хаусмана
2. Тест LM Бройша–Пагана
3. Тест AR(2)
4. Тест Вальда

**Поле для ответа:**

--	--	--	--

### Задание 10

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Что из перечисленного является примером пространственной матрицы весов?



1. Таблица сопряжённости между категориальными переменными
2. Матрица корреляций между переменными в модели
3. Матрица смежности, показывающая наличие границ между наблюдениями
4. Матрица рангов переменных по географическому положению

**Поле для ответа:**

--	--	--	--

### Задание 11

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой из следующих индексов используется для оценки степени глобальной пространственной автокорреляции переменной?

1. Индекс Джини
2. Индекс Морана
3. Индекс Херфиндаля
4. Индекс Вальда

**Поле для ответа:**

--	--	--	--

### Задание 12

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какое допущение лежит в основе стандартной пуассоновской модели для счётных данных?

1. Среднее значение и дисперсия случайной величины равны
2. Остатки модели нормально распределены
3. Ошибки модели независимы и гомоскедастичны
4. Наблюдаемые значения принимают только два возможных значения

**Поле для ответа:**

--	--	--	--

### Задание 13

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какая из приведённых метрик непосредственно отражает точность классификации в пробит-модели?

1. Коэффициент детерминации  $R^2$
2. Точность классификации (accuracy)
3. Средняя абсолютная ошибка (MAE)
4. Доля пропущенных значений

**Поле для ответа:**

--	--	--	--

**Задание 14**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Что означает термин «правостороннее цензурирование» в моделях длительности?

1. Значение переменной измерено с ошибкой
2. Длительность события зафиксирована точно
3. Время наступления события неизвестно, но превышает наблюдаемый период
4. Событие произошло до начала наблюдения

**Поле для ответа:**

--	--	--	--

**Задание 15**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Что означает "цензурирование" в модели анализа длительности (survival analysis)?

1. Исключение пропущенных наблюдений
2. Взвешивание выборки
3. Недоступность точного времени события у части наблюдений
4. Выравнивание шкалы времени

**Поле для ответа:**

--	--	--	--

**Задание 16**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какая из моделей позволяет учитывать как наличие события, так и его интенсивность в данных с большим числом нулей?

1. Пуассоновская модель
2. Отрицательная биномиальная модель
3. Zero-inflated модель
4. Модель множественного выбора

**Поле для ответа:**

--	--	--	--

**Задание 17**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какое допущение является ключевым для корректной работы упорядоченной пробит-модели?

1. Линейность связи между переменными
2. Нормальность распределения ошибок
3. Гомоскедастичность остатков
4. Независимость наблюдений

**Поле для ответа:**

--	--	--	--

**Задание 18**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой показатель используется для проверки избыточности дисперсии в пуассоновской регрессии?

1. Коэффициент детерминации  $R^2 < 1$
2. Отношение дисперсии к среднему  $> 1$
3. Сумма коэффициентов модели превышает 1
4. Отрицательный логарифм правдоподобия

**Поле для ответа:**

--	--	--	--

### Задание 19

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой тип модели следует использовать, если зависимая переменная представляет собой выбор из трёх и более неупорядоченных категорий?

1. Линейная регрессия
2. Упорядоченная пробит-модель
3. Мультиномиальная логит-модель
4. Пуассоновская регрессия

**Поле для ответа:**

--	--	--	--

### Задание 20

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Как интерпретируется коэффициент в пуассоновской регрессии?

1. Как изменение вероятности события
2. Как изменение логарифма ожидаемого значения зависимой переменной
3. Как предельный эффект
4. Как отношение шансов

**Поле для ответа:**

--	--	--	--

### Задание 21

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой тип данных характеризуется тем, что точное время события неизвестно, но известно, что оно не произошло до определённого момента?

1. Панельные данные
2. Цензурированные данные
3. Пропущенные данные
4. Агрегированные данные

**Поле для ответа:**

--	--	--	--

**Задание 22**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какое допущение отличает пуассоновскую модель от отрицательной биномиальной при моделировании счётных данных?

1. Гомоскедастичность ошибок
2. Равенство среднего и дисперсии
3. Нормальность распределения ошибок
4. Независимость наблюдений

**Поле для ответа:**

--	--	--	--

**Задание 23**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой статистический критерий чаще всего применяется для выбора между пуассоновской и отрицательной биномиальной моделями?

1. t-критерий
2. Критерий Акаике (AIC)
3. Тест множителя Лагранжа (LM)
4. Тест на нормальность остатков

**Поле для ответа:**

--	--	--	--

**Задание 24**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите один правильный вариант из предложенных и отметьте его в соответствующей клетке поля для ответа.

**Формулировка задания:**

Какой из следующих признаков указывает на избыточность дисперсии в счётных данных?

1. Среднее и дисперсия примерно равны

2. Дисперсия существенно превышает среднее
3. Отрицательные значения зависимой переменной
4. Много пропущенных значений

**Поле для ответа:**

--	--	--	--

### **Задания повышенного уровня сложности**

#### **Задание 1**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие признаки могут указывать на ограниченность модели с фиксированными эффектами (FE) или потенциальные проблемы её интерпретации?

**Варианты ответа:**

1.  $R^2$ -between близок к нулю, несмотря на высокий  $R^2$ -within
2. В модели не учтена гетероскедастичность ошибок
3. Результаты теста Хаусмана отвергают модель RE в пользу FE
4. Индивидуальные эффекты статистически значимы на уровне 5%
5. Все объясняющие переменные имеют ненулевую внутригрупповую вариацию

**Поле для ответа:**

--	--	--	--	--

#### **Задание 2**

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие из следующих утверждений описывают **типичные ошибки** при интерпретации результатов модели с фиксированными эффектами (FE)?

1. Исследователь считает, что модель с фиксированными эффектами позволяет оценить влияние пола и других неизменных характеристик
2. Исследователь утверждает, что FE-модель показывает различия между группами, а не внутри групп
3. Исследователь считает, что модель с фиксированными эффектами корректна только при равных количествах наблюдений в каждой группе
4. Исследователь использует кластеризацию по индивидам при оценивании стандартных ошибок.
5. Исследователь считает, что FE-модель устраняет смещение, вызванное пропущенными переменными, даже если они меняются во времени.

Поле для ответа:

--	--	--	--	--

### Задание 3

Компетенция: ПК-4

Тип задания: закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Результаты LM-теста Бреуша–Пагана (на наличие случайных эффектов) показали, что нулевая гипотеза **отвергается** на уровне значимости 5%. Какие выводы из этого следуют?

**Варианты ответа:**

1. Модель со случайными эффектами предпочтительнее объединённой регрессии
2. Наличие межгрупповой гетерогенности подтверждено
3. Следует обязательно использовать модель с фиксированными эффектами
4. Тест указывает на гетероскедастичность, и следует скорректировать стандартные ошибки
5. Можно использовать pooled OLS, так как межгрупповая гетерогенность отсутствует

Поле для ответа:

--	--	--	--	--

### Задание 4

Компетенция: ПК-4

Тип задания: закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

На рисунках представлены выводы команд для теста Breusch–Pagan LM в R и в Stata:

Lagrange Multiplier Test - (Breusch-Pagan) for random effects

```
data: ln_wage ~ age + I(age^2) + ttl_exp + I(ttl_exp^2) + tenure + I(tenure^2) + not_sms
chisq = 14779.98, df = 1, p-value < 2.2e-16
alternative hypothesis: significant effects
```

Breusch and Pagan Lagrangian multiplier test for random effects

```
ln_wage[idcode,t] = Xb + u[idcode] + e[idcode,t]
```

Estimated results:

	Var	sd = sqrt(Var)
ln_wage	.2283326	.4778416
e	.0845002	.2906892
u	.0665151	.2579053

Test: Var(u) = 0

```
chibar2(01) = 14779.98
Prob > chibar2 = 0.0000
```

Выберите все верные утверждения:

1. Случайные эффекты статистически значимы

2. Модель случайных эффектов более предпочтительна, чем объединённая модель
3. Модель фиксированных эффектов является более предпочтительной, чем модель случайных эффектов
4. Объединённая модель является более предпочтительной, чем модель случайных эффектов
5. Значение p-value указывает на значимое отличие модели случайных эффектов от pooled OLS

Поле для ответа:

--	--	--	--	--

## Задание 5

Компетенция: ПК-4

Тип задания: закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

На рисунках представлены выводы команд для теста Хаусмана в R и в Stata:

Hausman Test

```
data: ln_wage ~ age + I(age^2) + ttl_exp + I(ttl_exp^2) + tenure + I(tenure^2) + not_smsa + sc
chisq = 149.43, df = 8, p-value < 2.2e-16
alternative hypothesis: one model is inconsistent
```

```
. hausman . random_effects
```

	Coefficients			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	.	random_eff=s	Difference	S.E.
age	.0359987	.0368059	-.0008073	.0013177
c.age#c.age	-.000723	-.0007133	-.68e-06	.0000184
ttl_exp	.0334668	.0290208	.0044459	.001711
c.ttl_exp#c.ttl_exp	.0002163	.0003049	-.0000886	.000053
tenure	.0357539	.0392519	-.003498	.0005797
c.tenure#c.tenure	-.0019701	-.0020035	.0000334	.0000373
not_smsa	-.0890108	-.1308252	.0418144	.0062745
south	-.0606309	-.0868922	.0262613	.0081345

b = consistent under Ho and Ha; obtained from xtreg  
 B = inconsistent under Ha, efficient under Ho; obtained from xtreg  
 Test: Ho: difference in coefficients not systematic  
 $\chi^2(8) = (b-B)'[(V_b-V_B)^{-1}](b-B)$   
 = 149.43  
 Prob>chi2 = 0.0000

Выберите все верные утверждения:

1. Модель фиксированных эффектов является более предпочтительной
2. Оценки в модели случайных эффектов эффективны и состоятельны
3. Существует систематическая разница между оценками RE и FE
4. Модель случайных эффектов даёт смещённые оценки
5. Модель случайных эффектов является более предпочтительной

Поле для ответа:

--	--	--	--	--

## Задание 6

Компетенция: ПК-4



**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие проверки спецификации применимы к модели Ареллано–Бонда?

1. Тест Харке-Бера на нормальность остатков
2. Тест Хансена на валидность инструментов
3. LM-тест на пространственную автокорреляцию
4. Тест на автокорреляцию остатков первого и второго порядка (AR(1) и AR(2))
5. Тест на стационарность ряда (например, Дики-Фуллер)

**Поле для ответа:**

--	--	--	--	--

### Задание 7

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Прочитайте описание ситуации. Укажите корректную интерпретацию результатов.

**Формулировка задания:**

После оценки динамической панели методом Difference-GMM были получены следующие результаты:

- AR(1): значим
- AR(2): незначим
- Тест Хансена: p-value = 0.82

Как можно интерпретировать эти результаты с точки зрения валидности модели и инструментов?

**Поле для ответа:** \_\_\_\_\_

### Задание 8

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие из следующих факторов могут привести к снижению надёжности и точности оценок в динамической панели, оцененной методом Ареллано–Бонда?

1. Использование чрезмерно большого числа инструментов
2. Малое количество временных наблюдений на каждого индивида
3. Эндогенность объясняющих переменных, включая лаги зависимой
4. Применение лагов зависимой переменной в качестве регрессоров
5. Корректировка стандартных ошибок по кластерам

**Поле для ответа:**

--	--	--	--

### Задание 9

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:**

Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие из следующих утверждений корректно описывают специфику SAR и SLX моделей?

1. SAR-модель включает пространственный лаг зависимой переменной и требует специальных методов оценки (MLE/GMM), так как OLS даёт смещённые оценки.
2. SLX-модель включает пространственные лаги объясняющих переменных и может быть корректно оценена с помощью OLS.
3. В SLX-модели используются пространственные лаги объясняющих переменных (WX), но не включается лаг зависимой переменной (Wy).
4. SAR и SLX модели становятся эквивалентными при использовании нормализованной матрицы весов.
5. SLX-модель можно представить как частный случай SAR-модели при определённых ограничениях.

**Поле для ответа:**

--	--	--	--

### Задание 10

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие из следующих утверждений о пространственной матрице весов и глобальном индексе Морана являются корректными?

1. Пространственная матрица весов может быть построена на основе географической смежности или расстояний между объектами.
2. После нормализации строк пространственной матрицы весов все элементы становятся равными нулю или единице.
3. Глобальный индекс Морана измеряет степень пространственной автокорреляции в переменной на всей территории.
4. Значение индекса Морана не зависит от выбора матрицы весов.
5. Значительное положительное значение глобального индекса Морана свидетельствует о кластеризации схожих значений в пространстве.

**Поле для ответа:**

--	--	--	--

### Задание 11

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:**

Изучите результат пространственного анализа и дайте краткое обоснование вывода.

**Формулировка задания:**

В результате расчёта глобального индекса Морана по переменной «уровень безработицы» получено значение  $I = 0.45$ ,  $p\text{-value} = 0.001$ .

Объясните, как интерпретировать этот результат.

**Поле для ответа:** \_\_\_\_\_

### Задание 12

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Проанализируйте результат тестов на пространственную зависимость и сделайте вывод о модели.

**Формулировка задания:**

Вы провели тесты множителей Лагранжа на наличие пространственной зависимости в модели уровня преступности. Результаты:

- LM-лаг:  $\chi^2 = 6.5$ ,  $p\text{-value} = 0.011$
- LM-ошибка:  $\chi^2 = 2.2$ ,  $p\text{-value} = 0.137$
- Robust LM-лаг:  $\chi^2 = 5.9$ ,  $p\text{-value} = 0.015$
- Robust LM-ошибка:  $\chi^2 = 1.0$ ,  $p\text{-value} = 0.310$

Какой тип пространственной модели следует предпочесть?

**Поле для ответа:** \_\_\_\_\_

### Задание 13

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Формулировка задания:**

Вы используете мультиномиальную логит-модель (MNL). Какие утверждения верны для такого класса моделей?

1. Коэффициенты задают эффект по сравнению с базовой альтернативой
2. Коэффициенты не имеют прямой интерпретации в вероятностях
3. Предполагается независимость от нерелевантных альтернатив (ПА)
4. Модель допускает корреляцию между альтернативами
5. Предельные эффекты одинаковы для всех наблюдений

**Поле для ответа:**

--	--	--	--

#### Задание 14

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие особенности отличают модель Пуассона с избыточными нулями (Zero-inflated Poisson, ZIP) от классической пуассоновской регрессии?

**Варианты ответа:**

1. Учитывает два механизма генерации нулей
2. Включает в себя отдельные компоненты для бинарного и счётного процессов
3. Применяется исключительно в малых выборках
4. Не может использоваться для моделирования положительных значений зависимой переменной
5. Подходит при наличии структурных нулей в выборке

**Поле для ответа:**

--	--	--	--

#### Задание 15

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие статистические методы применяются при оценке длительности события в рамках непараметрических подходов?

**Варианты ответа:**

1. Оценка функции выживания по Каплану—Мейеру
2. Кумулятивная функция риска по Нельсону—Аалену
3. Двухшаговая оценка Ареллано-Бонда
4. Непараметрическое сравнение групп с помощью log-rank теста
5. Тест множителей Лагранжа (LM) на ограничение в виде равенства групп

**Поле для ответа:**

--	--	--	--

#### Задание 16

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите все правильные варианты ответов и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Какие показатели указывают на наличие избыточной дисперсии в модели счётных данных?

**Варианты ответа:**

1. Значительное превышение дисперсии над средним значением
2. Отрицательные значения зависимой переменной
3. Значение параметра дисперсии ( $\alpha$ ) в отрицательной биномиальной модели, статистически отличное от нуля
4. Значение логарифма правдоподобия выше, чем в пуассоновской модели
5. Наличие отрицательных остатков в модели

**Поле для ответа:**

--	--	--	--

### Задание 17

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите все корректные утверждения и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

Для оценки эффекта программы переподготовки (treatment = 1 – участвовал, = 0 – не участвовал) на частоту посещений центров занятости (count). Использована отрицательная биномиальная регрессия. Получены результаты

- при переменной training\_program:  $\beta = 0.6$  ( $p = 0.003$ ,  $SE = 0.2$ )
- LR-тест на избыточность дисперсии:  $\alpha = 2.1$  ( $p < 0.001$ )

Какие из утверждений корректно отражают результаты анализа?

1.  $IRR = \exp(0.6) \approx 1.82$  (95% ДИ: 1.23–2.71) указывает на увеличение ожидаемого числа визитов на 82% при участии в программе
2.  $\alpha > 0$  и результат LR-теста подтверждают наличие избыточной дисперсии, что обосновывает применение отрицательной биномиальной модели
3. Коэффициент 0.6 указывает на прирост вероятности визита в центр занятости на 0.6.
4. Для интерпретации в вероятностях требуется расчет предельных эффектов
5.  $\alpha > 0$  и результат LR-теста свидетельствуют против наличия избыточной дисперсии, что делает применение отрицательной биномиальной модели некорректным

**Поле для ответа:**

--	--	--	--	--

### Задание 18

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:** Выберите ВСЕ корректные утверждения о модели Пуассона с избыточными нулями (Zero-inflated Poisson, ZIP) модели и отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

При анализе количества заявок на пособие по безработице (переменная — count) была использована ZIP-модель и обычная пуассоновская регрессия (Poisson).

Полученные результаты:

- Коэффициент логит-компоненты ZIP:  $\beta = -1.2$  ( $p < 0.05$ )

- Коэффициент пуассоновской части ZIP:  $\beta = 0.4$  ( $p = 0.01$ )
- AIC (ZIP) = 420.7
- AIC (Poisson) = 446.3
- BIC (ZIP) = 440.1
- BIC (Poisson) = 463.5

**Какие из следующих утверждений корректны?**

1. Логит-компонента модели оценивает вероятность того, что наблюдение принадлежит к «структурным нулям».
2. Пуассоновская часть модели описывает только ненулевые наблюдения.
3. Более низкие значения AIC и BIC у ZIP-модели указывают, что она предпочтительнее стандартной пуассоновской.
4. ZIP-модель требует, чтобы все нулевые наблюдения были удалены перед оценкой.
5. ZIP-модель способна автоматически учитывать избыточную дисперсию, поэтому проверка на избыточную дисперсию не требуется.

**Поле для ответа:**

--	--	--	--	--

#### Задание 19

**Компетенция:** ПК-4

**Тип задания:** закрытый

**Инструкция:**

Проанализируйте результаты мультиномиальной логит-модели и выберите все корректные выводы, которые могут быть использованы для обоснования управленческих решений в транспортной политике. Отметьте их в соответствующих клетках поля для ответа.

**Формулировка задания:**

При моделировании выбора вида транспорта (базовая категория — «личный автомобиль») получены следующие результаты:

- Коэффициент при переменной *расстояние* (в км) для альтернативы «метро»:  $-1.0$  ( $p < 0.05$ )
- IRR (отношение шансов) для этой переменной: 0.37

Какие выводы можно сделать на основе этих результатов?

1. Коэффициент  $-1.0$  означает, что вероятность выбора метро снижается линейно на 1% при каждом километре.
2. Модель предполагает, что общее число поездок сокращается с увеличением расстояния, независимо от вида транспорта.
3. Увеличение расстояния поездки на 1 км статистически значимо снижает логарифм отношения шансов выбора метро вместо автомобиля на 1 единицу.
4. Шансы выбора метро относительно автомобиля снижаются примерно на 63% при увеличении расстояния на 1 км.
5. Для более длинных поездок разумно ожидать снижения доли пользователей метро по сравнению с автомобилем.

**Поле для ответа:**

--	--	--	--	--

### Задание 20

Компетенция: ПК-4

**Тип задания:** открытый

**Инструкция:** На основании сравнения двух моделей сделайте обоснованный вывод о выборе оптимальной модели для прогнозирования текучести кадров.

**Формулировка задания:**

Сравниваются две логит-модели прогнозирования увольнения сотрудников:

Модель 1 (базовая):

- Логарифм правдоподобия (LogL): -120
- AIC: 250

Модель 2 (с дополнительным фактором «удовлетворенность зарплатой»):

- LogL: -115
- AIC: 240

Дайте интерпретацию значений критериев. Обоснуйте выбор модели.

**Поле для ответа:** \_\_\_\_\_

### Задания высокого уровня сложности

#### Задание 1.

Компетенция: ПК-4

**Тип задания:** открытый

**Инструкция:** На основе описания задачи определите корректный подход к статистическому анализу.

**Формулировка задания:**

Вы получили панельные данные по регионам России с уровнем занятости  $u_i$ , ВРП на душу населения  $x_i$  и постоянными региональными характеристиками  $z_i$  за 2010–2020 гг. Вы предполагаете, что  $z_i$  коррелированы с  $x_i$ . Какие модели и процедуры позволяют учесть такую предполагаемую зависимость при оценке?

**Поле для ответа:** \_\_\_\_\_

#### Задание 2

Компетенция: ПК-4

**Тип задания:** открытый

**Инструкция:** Объясните, как интерпретировать результат теста и какое решение он предполагает.

**Формулировка задания:**

Вы провели тест Хаусмана и получили  $p\text{-value} = 0.87$ . Какой вывод вы сделаете и какую модель предпочтёте в дальнейшем анализе?

**Поле для ответа:** \_\_\_\_\_

### Задание 3

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Определите, какие меры следует предпринять для корректной оценки стандартных ошибок.

**Формулировка задания:**

Вы построили модель фиксированных эффектов на основе панельных данных по предприятиям. После предварительного анализа вы подозреваете, что ошибки автокоррелированы и гетероскедастичны. Что вы должны сделать перед интерпретацией значимости коэффициентов?

**Поле для ответа:** \_\_\_\_\_

### Задание 4

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** На основе описания интерпретируйте результат теста и сделайте обоснованный вывод о модели.

**Формулировка задания:**

Вы провели тест Бреуша–Пагана (LM) для сравнения модели случайных эффектов с объединённой регрессией (pooled OLS). Полученное значение  $p\text{-value} = 0.000$ . Какой вывод следует сделать?

**Поле для ответа:** \_\_\_\_\_

### Задание 5

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Ответьте на вопрос, касающийся условий применимости модели случайных эффектов. Объясните, почему соответствующее предположение критически важно, и какие последствия возникают при его нарушении. Ответ должен быть кратким (1–3 предложения), логически завершённым и обоснованным.

**Формулировка задания:**

В модели случайных эффектов (RE) одним из ключевых предположений является некоррелированность индивидуальных эффектов с объясняющими переменными. Почему это предположение важно, и что произойдёт, если оно нарушено?

**Поле для ответа:** \_\_\_\_\_

### Задание 6

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:**

Прочитайте описание модели и интерпретируйте ограничение, наложенное выбранным методом оценивания. Объясните, как оно влияет на спецификацию и интерпретацию результатов.

**Формулировка задания:**



Вы оценили модель фиксированных эффектов с использованием эффекта времени и эффекта региона. Модель включала как  $year_t$ , так и  $region_i$  фиктивные переменные. Однако при оценке модели статистический пакет вывел предупреждение об отсутствии идентификации из-за «совершенной мультиколлинеарности». Объясните, что это значит и как следует модифицировать модель.

**Поле для ответа:** \_\_\_\_\_

### Задание 7

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Внимательно прочитайте описание ситуации и ответьте на все вопросы в свободной форме. В ответе необходимо указать, какие тесты и процедуры следует провести, что именно они проверяют, и как интерпретировать возможные результаты. Ответ должен быть логически последовательным и продемонстрировать понимание методов анализа данных в динамических панелях.

**Формулировка задания:**

Вы оцениваете модель производительности труда на основе панельных данных по фирмам за 10 лет. Зависимая переменная — логарифм производительности. Среди объясняющих переменных — лаги зависимой переменной, инвестиции и численность персонала. Для устранения фиксированных фирменных эффектов и смещения, вызванного включением лагов зависимой переменной, вы используете обобщённый метод моментов (GMM) в первых разностях, то есть спецификацию модели Ареллано–Бонда.

После оценки вы получили следующие результаты:

- $p$ -значение теста на автокорреляцию остатков первого порядка ( $AR(1)$ )  $< 0.001$
- $p$ -значение теста на автокорреляцию второго порядка ( $AR(2)$ )  $= 0.28$
- $p$ -значение теста Хансена на валидность инструментов  $= 0.150$

**Вопросы:**

1. Какова цель тестов  $AR(1)$  и  $AR(2)$  в модели Ареллано–Бонда?
2. Как следует интерпретировать полученные значения этих тестов?
3. Что проверяет тест Хансена, и что означает его результат в данной ситуации?
4. Можно ли считать спецификацию модели корректной? Обоснуйте.

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.
- 3.
- 4.

### Задание 8

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Внимательно прочитайте описание. Ответьте в свободной форме. В ответе необходимо указать подход к построению модели, обоснование выбора инструментов и способы проверки корректности спецификации.

**Формулировка задания:**

Исследуется влияние прошлого экспорта на текущие инвестиции компаний. Используются панельные данные: 150 фирм, 5 лет. Предполагается наличие фиксированных эффектов и динамической зависимости.

Вы выбрали модель Ареллано–Бонда.

**Вопросы:**

1. Какие переменные следует инструментировать и почему?
2. Какие тесты позволят убедиться в корректности оценки?
3. Что означает, если тест Хансена отвергает нулевую гипотезу?

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.
- 3.

**Задание 9**

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Внимательно прочитайте ситуацию и ответьте в свободной форме. Укажите обоснование проблемы и подходы к её диагностике.

**Формулировка задания:**

В модели Ареллано–Бонда использовано 40 инструментов при числе наблюдений 150. Тест Хансена не отвергается ( $p = 0.80$ ). Однако при сокращении набора инструментов до 20 значение теста резко снижается.

**Вопросы:**

1. Что может означать высокая чувствительность теста Хансена к числу инструментов?
2. Почему важно ограничивать число инструментов?

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.

**Задание 10**

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Прочитайте описание. Ответьте на вопросы в свободной форме, опираясь на специфику данных и моделей.

**Формулировка задания:**

Исследуются эффекты кредитования на прибыльность малого бизнеса в динамике. Используются панельные данные: 80 фирм, 6 лет. После оценки динамической модели методом GMM выявлены значимые лаги прибыли и кредита. Тест AR(2) не отвергается, но  $H_0$  теста Хансена отвергается ( $p = 0.02$ ).

**Вопросы:**

1. Можно ли доверять оценкам?
2. Как можно улучшить спецификацию?

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.

### Задание 11

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Прочитайте описание ситуации. Ответьте на вопросы по интерпретации результатов и диагностике модели.

**Формулировка задания:**

Вы оцениваете стандартную линейную регрессию методом наименьших квадратов. По итогам оценки:

- все коэффициенты при регрессорах значимы и интерпретируемы;
- глобальный индекс Морана, рассчитанный для остатков модели, оказался статистически значимым ( $p < 0.01$ );
- LM-тест на пространственную автокорреляцию ошибок также показал значимость ( $p < 0.05$ ).

**Вопросы:**

3. Какая модель пространственной регрессии предпочтительна и почему?
4. Как проверить, устранена ли пространственная зависимость после оценки модели?

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.

### Задание 12

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:**

Проанализируйте результат теста и предложите следующий шаг. Обоснуйте необходимость выбранной методологии.

**Формулировка задания:**

Вы провели тест глобального индекса Морана для переменной «уровень бедности» в 60 регионах. Получено значение  $I = 0.31$ ,  $p\text{-value} = 0.006$ . На тематической карте визуально наблюдается кластеризация бедности в нескольких соседних регионах.

Какие выводы можно сделать на основе теста и карты? Какие модели и методы стоит применить для дальнейшего анализа?

**Поле для ответа:** \_\_\_\_\_

### Задание 13

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Внимательно прочитайте описание ситуации и ответьте в свободной форме. В ответе необходимо обосновать выбор модели, описать её структуру и предложить процедуры для проверки качества и пригодности модели с точки зрения обработки статистических данных.

**Формулировка задания:**

Вы анализируете данные об использовании поликлинических услуг: переменная — количество визитов пациента за год. В выборке большая доля наблюдений равна нулю

(пациенты не обращались вовсе), много значений 1–2, но также встречаются значения 10 и более визитов. Оценка пуассоновской регрессии показывает, что дисперсия существенно превышает среднее значение переменной, а также фиксируется избыток нулей по сравнению с предсказаниями модели.

**Вопросы:**

3. Почему пуассоновская модель может быть некорректна в данной ситуации?
4. Какую альтернативную спецификацию модели целесообразно использовать?

**Поле для ответа:** \_\_\_\_\_

- 1.
- 2.

**Задание 14**

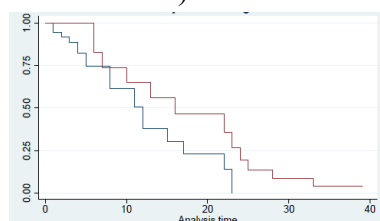
**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Ответьте на вопрос в свободной форме, кратко объяснив вывод.

**Задание:**

На графике представлены функции выживания заемщиков — вероятность «дожить» до определённого момента времени без дефолта. Сравниваются две группы: участвовавшие в программе реструктуризации ( $\text{posttran} = 1$ , красная линия) и не участвовавшие ( $\text{posttran} = 0$ , синяя линия).



```
. sts test posttran
```

Equality of survivor functions  
Log-rank test

posttran	Observed events	Expected events
0	30	31.20
1	45	43.80
Total	75	75.00

chi2(1) = 0.13  
Pr>chi2 = 0.7225

**Вопрос:**

Можно ли сделать вывод о статистически значимом влиянии программы реструктуризации на время до дефолта? Дайте обоснованный ответ, интерпретируя визуальные данные и результаты теста.

**Поле для ответа:** \_\_\_\_\_

**Задание 15**

**Компетенция:** ПК-4

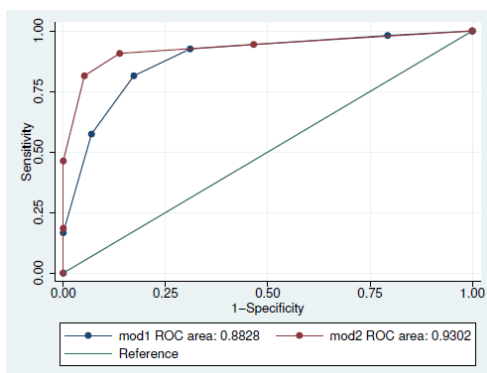
**Тип задания:** открытый

**Инструкция:**

На основе ROC-графика сравните предсказательную способность двух логистических моделей.

**Формулировка задания:**

На графике представлены ROC-кривые двух логистических моделей,  $\text{mod1}$  и  $\text{mod2}$ , для предсказания одного и того же бинарного исхода.



### Вопрос:

А. Какая модель предсказывает исход точнее? Обоснуйте по графику и AUC.

В. Какой смысл имеет линия Reference? Что можно сказать о модели, если её ROC-кривая близка к этой линии?

Поле для ответа: \_\_\_\_\_

А.

В.

### Задание 16

Компетенция: ПК-4

Тип задания: открытый

Инструкция: Ответьте на вопрос в свободной форме, кратко объяснив вывод.

Формулировка задания:

В логистической модели получена следующая классификационная таблица:

	Истинное значение $Y = 1$	Истинное значение $Y = 0$	Всего
Предсказано $Y = 1 (+)$	20	10	30
Предсказано $Y = 0 (-)$	5	15	20
Всего	25	25	50

На основании таблицы рассчитайте:

1. Чувствительность (Sensitivity, True Positive Rate)
2. Специфичность (Specificity, True Negative Rate)
3. Долю правильно классифицированных наблюдений (Accuracy)

Поле для ответа: \_\_\_\_\_

1.

2.

3.

### Задание 17

Компетенция: ПК-4

Тип задания: открытый

Инструкция:.

Вопрос:

Вы оцениваете модель выбора транспортного средства: автомобиль, автобус, велосипед.

После включения новой альтернативы (метро), предсказанные вероятности по оставшимся трём вариантам значительно изменились, особенно между автомобилем и автобусом.

Какой вывод можно сделать?

Поле для ответа: \_\_\_\_\_

### Задание 18

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задачи. Найдите аналитическое выражение для предельного эффекта. Ответ должен быть выражен через параметры модели.

**Формулировка задания:**

Исследуется вероятность участия в добровольной медицинской страховке. Модель задана в виде пробит-спецификации:

$$\Pr(Y_i = 1 | age_i, Z_i) = \Phi(\beta_0 + \beta_1 \times age_i + \beta_2 \times age_i^2 + Z_i' \gamma)$$

где  $\Phi(\cdot)$  — функция стандартного нормального распределения,  $age_i$  — возраст,  $Z_i$  — вектор других переменных.

**Вопрос:**

1. Запишите выражение для предельного эффекта возраста: через параметры модели и значение возраста.
2. Как интерпретировать знак этого эффекта?

Поле для ответа: \_\_\_\_\_

1.

2.

### Задание 19

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задачи. Найдите аналитическое выражение для предельного эффекта. Ответ должен быть выражен через параметры модели.

**Формулировка задания:**

Исследуется вероятность того, что человек имеет сберегательный вклад. Используется пробит-модель:

$$\Pr(Y_i = 1 | X_i) = \Phi(\beta_0 + \beta_1 \cdot age_i + \beta_2 \cdot female_i + \beta_3 \cdot age_i \cdot female_i),$$

где  $female_i = 1$  если респондент — женщина, и 0 — если мужчина.

**Вопросы:**

1. Выведите предельный эффект возраста для мужчин и женщин (по отдельности).
2. Как интерпретировать коэффициент  $\beta_3$  при взаимодействии  $age_i \cdot female_i$ ?

Поле для ответа: \_\_\_\_\_

1.

2.

### Задание 20

**Компетенция:** ПК-4

**Тип задания:** открытый

**Инструкция:** Прочитайте описание задачи. Ответьте на вопросы по спецификации модели и корректной интерпретации её параметров.

### Формулировка задания:

Вы анализируете данные о занятости женщин в зависимости от наличия детей. В качестве зависимой переменной используется бинарный индикатор *employed*, принимающий значение 1, если женщина работает, и 0 — иначе.

В регрессию включены переменные:

- *kids* — количество детей;
- *educ* — количество лет образования;
- *married* — бинарная переменная, равная 1, если женщина замужем;
- *age* — возраст;
- взаимодействие переменных  $kids \times married$ .

Оценка логит-модели дала следующие результаты (все коэффициенты статистически значимы):

- $\beta_1$  (при переменной *kids*) =  $-0.35$
- $\beta_2$  (при переменной *educ*) =  $0.20$
- $\beta_3$  (при переменной *married*) =  $0.10$
- $\beta_4$  (при переменной *age*) =  $0.01$
- $\beta_5$  (при взаимодействии переменных  $kids \times married$ ) =  $0.30$

### Вопросы:

1. Как интерпретировать знак и смысл взаимодействия  $kids \times married$  в логит-модели?
2. Учитывая оценки коэффициентов, у кого выше вероятность занятости: у замужней женщины с 2 детьми или у незамужней женщины с 2 детьми (при прочих равных)? Обоснуйте.

Поле для ответа: \_\_\_\_\_

- 1.
- 2.

## 5 Средства оценки индикаторов достижения компетенций

Таблица 5

**Средства оценки индикаторов достижения компетенций**

Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Средства оценки (в соот. с Таблицами 5, 7)
ПК-3	ИД.ПК-3.1. ИД.ПК-3.2.	домашнее задание (ДЗ), тестирование
ПК-4	ИД.ПК-4.1. ИД.ПК-4.2. ИД.ПК-4.3. ИД.ПК-4.4.	домашнее задание (ДЗ), тестирование

Таблица 6

**Описание средств оценки индикаторов достижения компетенций**

Средства оценки (в соот. с Таблицами 5, 7)	Рекомендованный план выполнения работы
Домашнее задание	Магистрант в ходе подготовки и выполнения домашнего задания по темам дисциплины, показывает способность совершать следующий набор профессиональных действий, получивший развитие в рамках данной дисциплины: — Составляет прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом на основании самостоятельно разработанных моделей с использованием математических,

Средства оценки (в соот. с Таблицами 5, 7)	Рекомендованный план выполнения работы
	<p>статистических и инструментальных методов, а также современных интеллектуальных технологий</p> <p>— Применяет математические методы и информационные технологии, проводит статистические расчеты, работает с различными источниками статистической информации, ведет аналитическую работу с данными, разрабатывает и совершенствует методологии сбора и обработки статистических данных</p>
Тестирование	<p>Магистрант в ходе подготовки и выполнения тестирования по темам дисциплины, показывает способность совершать следующий набор профессиональных действий, получивший развитие в рамках данной дисциплины:</p> <p>— Составляет прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом на основании самостоятельно разработанных моделей с использованием математических, статистических и инструментальных методов, а также современных интеллектуальных технологий</p> <p>— Применяет математические методы и информационные технологии, проводит статистические расчеты, работает с различными источниками статистической информации, ведет аналитическую работу с данными, разрабатывает и совершенствует методологии сбора и обработки статистических данных</p>