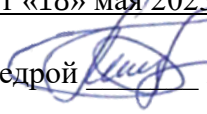


Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Наумова Наталия Александровна
Должность: Ректор
Дата подписания: 24.10.2024 14:21:41
Уникальный программный ключ:
6b5279da4e034bffa79172803da5b7b559f669e2

МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение высшего образования
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ПРОСВЕЩЕНИЯ
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ПРОСВЕЩЕНИЯ)

Кафедра географии, геоэкологии и природопользования

УТВЕРЖДЕН
на заседании кафедры
Протокол от «18» мая 2023 г. № 11

И.о.зав.кафедрой  Крылов П.М.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

по дисциплине

Методы математической обработки данных

Направление подготовки

44.03.05 Педагогическое образование с двумя профилями подготовки

Профиль

География и экономическое образование

Мытищи
2023

СОДЕРЖАНИЕ

	Стр.
Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы	3
Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания	3
Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы	6
Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций	13

УП 2023 г.г. набора

1.Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы

Код и наименование компетенции	Этапы формирования компетенций
УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	1.Работа на учебных занятиях 2.Самостоятельная работа
ОПК-9. Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	1.Работа на учебных занятиях 2.Самостоятельная работа

2.Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Оцениваемые компетенции	Уровень	Этап формирования	Описание показателей	Критерии оценивания	Шкала оценивания
УК-1.	пороговый	1.Работа на учебных занятиях 2.Самостоятельная работа	Знает и понимает: методики поиска, сбора и обработки информации; актуальные российские и зарубежные источники информации в сфере профессиональной деятельности; метод системного анализа Умеет: применять методики поиска, сбора и обработки информации; осуществляет критический анализ и синтез информации, полученной из разных	Конспект, контрольная работа, реферат, проект, лабораторная работа	Шкала оценивания конспекта Шкала оценивания контрольной работы Шкала оценивания реферата Шкала оценивания проекта Шкала оценивания лабораторной работы

			источников; применяет системный подход для решения поставленных задач		
	продвину- тый	1.Работан на учебных занятиях 2.Самостоятел- ьная работа	Знает и понимает: методики поиска, сбора и обработки информации; актуальные русские и зарубежные источники информации в сфере профессиональ- ной деятельности; метод системного анализа Умеет: применять методики поиска, сбора и обработки информации; осуществляет критический анализ и синтез информации, полученной из разных источников; применяет системный подход для решения поставленных задач Владеет: методами поиска, сбора и обработки, критического анализа и	Конспект, контрольн- ая работа, реферат, проект, лаборатор- ная работа	Шкала оценивани- я конспекта Шкала оценивани- я контрольн- ой работы Шкала оценивани- я реферата Шкала оценивани- я проекта Шкала оценивани- я лаборатор- ной работы

			синтеза информации; методикой системного подхода для решения поставленных задач		
ОПК-9.	пороговый	1.Работан на учебных занятиях 2.Самостоятельная работа	Знает и понимает: современные программные средства и цифровые ресурсы для решения исследовательских задач Умеет: применять современные программные средства и цифровые ресурсы для решения исследовательских задач (стандартные статистические пакеты для осуществления обработки полученных эмпирических и экспериментальных данных)	Конспект, контрольная работа, реферат, проект, лабораторная работа	Шкала оценивания конспекта Шкала оценивания контрольной работы Шкала оценивания реферата Шкала оценивания проекта Шкала оценивания лабораторной работы
	продвинутый	1.Работан на учебных занятиях 2.Самостоятельная работа	Знает и понимает: современные программные средства и цифровые ресурсы для решения исследовательских задач Умеет: применять современные	Конспект, контрольная работа, реферат, проект, лабораторная работа	Шкала оценивания конспекта Шкала оценивания контрольной работы Шкала оценивания реферата Шкала

			программные средства и цифровые ресурсы для решения исследовательских задач (стандартные статистические пакеты для осуществления обработки полученных эмпирических и экспериментальных данных) Владеет (навыками и/или опытом деятельности): навыками применения современных программных средств и цифровых ресурсов для решения исследовательских задач (стандартных статистических пакетов для осуществления обработки полученных эмпирических и экспериментальных данных)		оценивание проекта Шкала оценивания лабораторной работы
--	--	--	---	--	--

Описание шкал оценивания

Шкала оценивания конспекта

Представление конспекта оценивается от 0 до 10 баллов. В качестве оценки используется следующие критерии:

8-10 баллов. В содержании конспекта соблюдена логика изложения вопроса темы; материал изложен в полном объеме; выделены ключевые моменты вопроса, материал изложен понятным языком; приведены примеры, иллюстрирующие ключевые моменты темы. (оценка отлично);

7-6 баллов. В содержании конспекта соблюден литературный стиль изложения,

прослеживается некоторая неясность и нечеткость изложения, иллюстрационные примеры приведены с замечаниями (оценка хорошо);

4-5 баллов. В содержании конспекта не соблюден литературный стиль изложения, прослеживается неясность и нечеткость изложения, иллюстрационные примеры приведены не в полном объеме (оценка удовлетворительно);

0–3 балла. Конспект составлен небрежно и неграмотно, имеются нарушения логики изложения материала темы, не приведены иллюстрационные примеры, не выделены ключевые моменты темы. (оценка неудовлетворительно).

Шкала оценивания контрольной работы

По результатам проверки контрольной работы выставляется балл от 0 до 10.

Если контрольная работа отвечает следующим критериям: работа написана студентом самостоятельно и ней в полном объеме раскрыты вопросы контрольных заданий; использована монографическая и специальная литература; работа содержит правильную формулировку понятий и категорий; в освещении вопросов заданий не содержится грубых ошибок; при решении задач сделаны правильные и аргументированные выводы – 4-10 баллов

Если студент не справился с заданиями, в работе не раскрыто основное содержание вопросов, имеются ошибки в решении задач и освещении вопросов заданий, а так же имеются явные признаки плагиата. Оформление работы не соответствует предъявляемым требованиям - 0-3 баллов

Работа, по результатам проверки которой выставлена оценка «0» возвращается студенту на доработку, причём, до тех пор пока студент не предоставит контрольную работу с доработанными недочётами и исправленными ошибками.

Шкала оценивания реферата

Критерии	Показатели
Новизна реферированного теста 5 балла	актуальность проблемы и темы; наличие авторской позиции, самостоятельность суждений
Степень раскрытия сущности проблемы 5 балла	соответствие содержания теме и плану реферата; умение работать с литературой, систематизировать и структурировать материал;
Обоснованность выбора источников 5 балл	круг, полнота использования литературных источников по проблеме
Соблюдение требований к оформлению 5 балл	правильное оформление ссылок на используемую литературу; соблюдение требований к оформлению и объему реферата
Грамотность 5 балл	отсутствие орфографических и синтаксических ошибок, стилистических погрешностей; литературный стиль.

Шкала оценивания проекта

Шкала оценивания	Критерии оценивания
Продвинутый	Студент владеет навыками письменной речи с использованием

уровень 8-10 баллов	правил логики, аргументации на высоком уровне. Научно-исследовательская работа носит самостоятельный характер и выполнена на достаточно высоком научном уровне. Материал работы основан на действующем законодательстве, научной и учебной литературе. Автор изучил и проанализировал материалы правоприменительной (судебной практики). Выявлены проблемы правового регулирования, практики применения законодательства и сформулированы предложения по их решения, в том числе законодательного характера. Использованы методы научного познания. Позиция автора обоснована и аргументирована, отражает его точку зрения относительно исследуемой проблемы (вопроса). Работа отвечает всем требованиям научной работы (условиям конкурса). Результаты научно-исследовательской работы правильно оформлены в виде научного доклада, сообщения (выступления), научной статьи (в зависимости от формы научно-исследовательской деятельности) - 8-10 баллов
Достаточный 5-7 баллов	Научно-исследовательская работа самостоятельно и основана на действующем законодательстве, научной и учебной литературе, изученной правоприменительной (судебной) практике. Использованы некоторые методы научного познания. Позиция автора обоснована, но не достаточно чётко аргументирована. Выявлены пробелы правового регулирования и сформулировано хотя бы одно предложение по их решению. Работа отвечает всем основным требованиям научной работы (условиям конкурса). Результаты научно-исследовательской работы правильно оформлены в виде научного доклада, сообщения (выступления), научной статьи (в зависимости от формы научно-исследовательской деятельности) - 5-7 баллов
Пороговый 1-4 балла	Научно-исследовательская работа имеет существенные недостатки и нуждается в доработке. В частности, недостаточно обоснована позиция автор, не указаны проблемы правового регулирования, недостаточно проанализированы материалы правоприменительной (судебной) практики; имеются недостатки в оформлении работы 1-4 баллов.

Шкала оценивания лабораторной работы

24-30 баллов: выполнение всех работ семестра без замечаний и посещение более 80% занятий семестра. Студент чётко и без ошибок ответил на все контрольные вопросы. Работа выполнена полностью. Студент владеет теоретическим материалом, отсутствуют ошибки при описании теории, формулирует собственные, самостоятельные, обоснованные, аргументированные суждения, представляет полные и развернутые ответы на дополнительные вопросы.

16-23 баллов: лабораторные работы семестра выполнены с небольшими замечаниями; студент ответил на контрольные вопросы, но с небольшими замечаниями. Работа выполнена полностью. Студент владеет теоретическим материалом, отсутствуют ошибки при описании теории, формулирует собственные, самостоятельные, обоснованные, аргументированные суждения, допуская незначительные ошибки при ответе на дополнительные вопросы.

8-15 баллов: задания работ выполнены с замечаниями; студент ответил на контрольные вопросы с замечаниями. Работа выполнена полностью. Студент владеет теоретическим материалом на минимально допустимом уровне, отсутствуют ошибки при описании теории, испытывает затруднения в формулировке собственных обоснованных и аргументированных суждений, допуская незначительные ошибки на дополнительные вопросы.

0-7 баллов: студент не выполнил или выполнил неправильно лабораторные работы; студент ответил на контрольные вопросы с серьезными ошибками или не ответил на контрольные вопросы. Студент не владеет теоретическим материалом, допускает грубые ошибки, испытывает затруднения в формулировке собственных суждений, неспособен ответить на дополнительные вопросы.

3.Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Текущий контроль	
УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	
Знает и понимает: методики поиска, сбора и обработки информации; актуальные российские и зарубежные источники информации в сфере профессиональной деятельности; метод системного анализа	
Перечень вопросов для устного опроса	
1. Статистический анализ причинно-следственных связей в водных экосистемах. 2. Статистические методы прогнозирования показателей водных экосистем. 3. Статистический анализ качества речных вод в промышленном регионе. 4. Статистический анализ качества и уровня жизни семей в современной России. 5. Статистический анализ жилищных условий различных типов домохозяйств и оценка эффективности жилищной политики. 6. Статистический анализ факторов рождаемости. 7. Статистический анализ уровня образования россиян: тенденции и дифференциация. 8. Статистическое исследование состояние здоровья населения. 9. Статистический анализ условий жизни домохозяйств в Российской Федерации. 10. Статистический анализ уровня счастья людей в Российской Федерации. 11. Статистический анализ досуговой активности россиян 12. Применение фильтра Калмана в задачах оценивания и прогнозирования 13. Взаимосвязь системы таблиц "затраты – выпуск" и классической модели "затраты – выпуск" В. Леонтьева 14. Статистический анализ региональной дифференциации по уровню инновационного развития экономики. 15. Статистическое исследование рынка жилья в регионах России.	
Умеет: применять методики поиска, сбора и обработки информации; осуществляет критический анализ и синтез информации, полученной из разных источников; применяет системный подход для решения поставленных задач	
Перечень вопросов для тестовых заданий	
• Какой метод используется для агрегации данных в базе данных? • Гипертекстовая передача данных Метод обратного вызова SQL-запросы Протокол	

передачи гипертекста

- Какой инструмент чаще всего применяется для визуализации данных?
- Программирование на C++ Графические библиотеки Алгоритмы сортировки Сетевые протоколы
- Что такое ETL в контексте обработки данных?
- Экстремально техническая литература Электротранспортная лаборатория Извлечение, трансформация и загрузка Экспериментальная теория логики
- Какой метод используется для заполнения пропущенных значений в данных?
- Цифровая модуляция Интерполяция Метод случайного выбора Дискретизация данных
- Что такое нормализация данных?
- Процесс создания резервных копий Процесс устранения избыточности данных и улучшения их структуры Процесс кодирования данных Процесс шифрования данных
- Какой метод используется для выявления аномалий в данных?
- Метод случайных чисел Анализ аномалий Метод определения центра масс Метод обратного распространения ошибки
- Что такое регуляризация в машинном обучении?
- Спецификация оборудования Метод добавления штрафа за сложность модели для предотвращения переобучения Технология виртуальной реальности Метод оценки времени выполнения программы
- Какой метод используется для уменьшения размерности данных?
- Метод повышения контрастности изображения Метод главных компонент Метод компьютерного зрения Метод квадратичного программирования
- Что такое кросс-валидация в машинном обучении?
- Метод расчета статистических ошибок Метод обработки изображений Метод разделения данных на обучающую и тестовую выборки для оценки производительности модели Метод обучения без учителя
- Какой метод применяется для поиска ассоциативных правил в данных?
- Метод обучения с подкреплением Алгоритм Ah2riori Метод многокритериальной оптимизации Метод генетических алгоритмов
- Какой метод используется для выделения наиболее важных признаков в данных?
- Анализ главных компонент Метод k-средних Линейная регрессия Деревья решений
- Что такое метод регуляризации в контексте машинного обучения?
- Процесс очистки данных от выбросов Метод устранения шума в изображениях Добавление штрафа за сложность модели Выбор оптимального числа кластеров
- Какой алгоритм используется для классификации объектов на две группы?
- Логистическая регрессия Случайный лес Метод опорных векторов k-ближайших соседей
- Что представляет собой метод бутстрапа в статистике?
- Метод аугментации данных Алгоритм обучения без учителя Создание выборок путем случайного выбора с повторениями Метод для определения статистической значимости
- Какой метод используется для заполнения пропущенных значений в данных?
- Кластеризация Регрессия Интерполяция Ансамблевые методы
- Что такое однородность в контексте кластерного анализа?
- Отсутствие выбросов Высокая степень разделения классов Схожесть объектов внутри одного кластера Минимизация функции потерь

- Какой метод используется для снижения размерности данных?
- Метод опорных векторов t-CNE Градиентный спуск Алгоритм k-средних
- Что представляют собой выбросы в данных?
- Экстремально высокие значения Малочисленные классы в задаче классификации Точки, не соответствующие общему тренду данных Аномалии или необычные значения, отклоняющиеся от общего распределения
- Какой метод используется для определения оптимального числа кластеров в кластерном анализе?
- Алгоритм k-средних Метод главных компонент Метод локтя Иерархическая кластеризация
- Что такое перекрестная проверка (кросс-валидация) в машинном обучении?
- Метод сжатия изображений Техника разделения данных на обучающую и тестовую выборки для оценки модели Алгоритм обучения с учителем Процесс нормализации данных
- Какой метод обработки данных используется для заполнения отсутствующих значений в наборе данных?
- А) Метод случайного выбора Б) Метод последовательного заполнения В) Метод исключения Г) Метод дублирования
- Что такое EDA в контексте обработки данных?
- А) Электронная цифровая аналитика Б) Анализ исследования данных В) Эффективная динамическая адаптация Г) Энергетическая декомпозиция алгоритма
- Какой из перечисленных методов используется для кодирования категориальных переменных?
- А) Анализ главных компонент Б) Метод k-ближайших соседей В) One-Hot Encoding Г) Регрессионный анализ

Владеет:

методами поиска, сбора и обработки, критического анализа и синтеза информации; методикой системного подхода для решения поставленных задач

Перечень тем для докладов

1. Регрессионный анализ – возможности и перспективы использования в экономических исследованиях.
2. Обработка географических данных с помощью дисперсионного анализа.
3. Обработка экономических данных с помощью дисперсионного анализа.
4. Дисперсионный анализ – возможности и перспективы использования в географических исследованиях.
5. Дисперсионный анализ – возможности и перспективы использования в экономических исследованиях.
6. Особенность использования математических методов в различных географических направлениях.
7. Особенность использования математических методов в отдельных областях экономики.
8. Каковы особенности сбора и первичной обработки информации?
9. Перечислите виды и назовите основные характеристики специализированных программ для статистической обработке географической информации.
10. Перечислите виды и назовите основные характеристики специализированных программ для статистической обработке экономических данных.
11. Мерные и счетные количественные признаки – в чем разница?
12. Виды средних величин?

13. Что характеризует среднее квадратическое отклонение? 14. О чем судят по критерию Стьюдента?
Промежуточная аттестация
УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач
Знает и понимает: методики поиска, сбора и обработки информации; актуальные российские и зарубежные источники информации в сфере профессиональной деятельности; метод системного анализа Умеет: применять методики поиска, сбора и обработки информации; осуществляет критический анализ и синтез информации, полученной из разных источников; применяет системный подход для решения поставленных задач Владеет: методами поиска, сбора и обработки, критического анализа и синтеза информации; методикой системного подхода для решения поставленных задач
Задания, необходимые для оценивания сформированности компетенции
Перечень вопросов для зачета
1. Распределение Хи-квадрат. Определить круг задач, в которых используется распределение Хи-квадрат. 2. Распределение Стьюдента, (t - распределение). Определить круг задач, в которых используется распределение Стьюдента. 3. Распределение Фишера, (F - распределение). Определить круг задач, в которых используется распределение Фишера. 4. Отсев грубых измерений/ 5. Гипотезы простые, сложные, односторонние, двусторонние. 6. Проверка простой гипотезы H_0 о величине математического ожидания нормального распределения против сложной гипотезы H_1 при неизвестной дисперсии. 7. Применение критерия отношения правдоподобия для проверки простой гипотезы H_0 о виде распределения. 8. Непараметрический критерий об однородности двух выборок/ 9. Непараметрический критерий для проверки гипотезы о некоррелированности двух распределений/ 10. Регрессионный анализ. Модель регрессии. Оценка параметров регрессии. 11. Регрессионный анализ. Анализ ошибок. Остаточная дисперсия. Дисперсия, обусловленная регрессией. 12. Регрессионный анализ. Коэффициент детерминации и мера неопределённости при построении модели регрессии. 13. Дисперсия оценок параметров регрессии.

Текущий контроль
ОПК-9. Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности
Знает и понимает: современные программные средства и цифровые ресурсы для решения исследовательских задач
Перечень вопросов для устного опроса

• Основные модельные распределения. устный опрос , примерные вопросы: Нормальное распределение, основные свойства и условия применимости. Центральная предельная теорема. • Проверка гипотез. устный опрос , примерные вопросы: Постановка задачи проверки гипотез. Требования, предъявляемые к критериям. • Критерий отношения правдоподобия домашнее задание , примерные вопросы: Распределение χ^2 и область его применения. Распределение Стьюдента, (t - распределение). Распределение Фишера, (F - распределение). • Непараметрические критерии. домашнее задание , примерные вопросы: Несмещённый критерий, несмещённая критическая область. Состоятельность критерия. Равномерно наиболее мощный критерий (РНМ-критерий). Отсев грубых измерений по критерию Стьюдента. • Регрессионный анализ. домашнее задание , примерные вопросы: Пример проверки сложных гипотез ? $H_0 : m = m_0, s^2$ - неизвестно, $H_1 : m \neq m_0, s^2$ - неизвестно. Отсев грубых измерений при малых выборках. Отсев грубых измерений при больших выборках. • Задача линейной регрессии. домашнее задание , примерные вопросы: Неравенство Рао-Крамера и предел эффективности оценок параметров. Оценка параметров для выбранного закона распределения. • Одномерная линейная регрессия. домашнее задание , примерные вопросы: Коэффициент детерминации. Проверка значимости коэффициента детерминации. Применение медианной фильтрации для устранения импульсных помех.
Умеет: применять современные программные средства и цифровые ресурсы для решения исследовательских задач (стандартные статистические пакеты для осуществления обработки полученных эмпирических и экспериментальных данных)
Перечень вопросов для тестовых заданий • Что представляет собой метод группировки (кластеризации) в обработке данных? • А) Прогнозирование временных рядов Б) Классификация объектов В) Объединение похожих объектов в группы Г) Оценка статистической значимости Какой алгоритм используется для уменьшения размерности данных путем отбора наиболее важных признаков? А) Метод опорных векторов Б) Отбор признаков с использованием случайного леса В) Градиентный бустинг Г) Алгоритм k-средних • Какие из перечисленных методов относятся к обучению без учителя? • А) Линейная регрессия Б) Кластеризация В) Метод опорных векторов Г) Понижение размерности • Что представляет собой метод стандартизации данных? А) Преобразование данных в бинарный формат Б) Преобразование данных в отрицательный диапазон В) Приведение данных к нулевому среднему и единичному стандартному отклонению Г) Преобразование данных в целочисленный формат • Какой метод используется для предсказания числовой переменной на основе линейной зависимости? • А) Метод k-ближайших соседей Б) Линейная регрессия В) Алгоритм дерева решений Г) Метод главных компонент • Какой метод обработки текстовых данных позволяет представить слова в виде векторов с учетом их семантического значения? • А) Метод TF-IDF Б) Метод ближайших соседей В) Word Embeddings Г) Метод биграмм

- Какой метод используется для оценки важности признаков в задачах классификации?
- А) Алгоритм k-средних Б) Важность признаков с использованием случайного леса В) Метод опорных векторов Г) Градиентный бустинг
- Какие методы обработки данных широко используются для устранения пропущенных значений?
- Метод агрегации Метод заполнения (imputation) Метод исключения Метод сортировки
- Какой метод обработки данных применяется для уменьшения размерности данных путем отбора наиболее значимых признаков?
- Метод нормализации Метод кластеризации Метод отбора признаков Метод регрессии
- Какой из перечисленных методов используется для выделения основных компонентов и снижения размерности данных?
- Метод фильтрации Метод биннинга Метод агрегации Метод главных компонент (h2CA)
- Какой метод обработки текстовых данных преобразует слова в числовые векторы?
- Метод кодирования One-Hot Метод векторизации (Word Embeddings) Метод стемминга Метод TF-IDF
- Какой метод обработки категориальных данных заменяет каждую категорию на среднее значение целевой переменной для этой категории?
- Метод биннинга Метод преобразования меток (Label Encoding) Метод частотного кодирования Метод среднего кодирования (Mean Encoding)
- Какой метод используется для обнаружения и удаления выбросов в данных?
- Метод нормализации Метод обнаружения выбросов (Outlier Detection) Метод агрегации Метод биннинга
- Какой метод преобразования данных используется для приведения значений признаков к заданному диапазону?
- Метод главных компонент (h2CA) Метод стандартизации Метод масштабирования (Scaling) Метод кластеризации
- Какой метод преобразования данных может использоваться для преобразования нелинейных зависимостей между признаками?
- Метод биннинга Метод стандартизации Метод полиномиальных признаков Метод кодирования One-Hot

Владеет (навыками и/или опытом деятельности):

навыками применения современных программных средств и цифровых ресурсов для решения исследовательских задач (стандартных статистических пакетов для осуществления обработки полученных эмпирических и экспериментальных данных)

Перечень тем для докладов

15. Учёные-математики – влияние на географические исследования.
16. Учёные-математики – влияние на эконометрику.
17. Учёные-географы - математические методы в географии.
18. Учёные-экономисты - математические методы в экономике.
19. Особенности написания научных статей с использованием математических методов.
20. Проблемы применения математических законов для географических исследований.
21. Проблемы применения математических законов для экономических исследований.
22. Возможности описательной статистики для анализа вариаций.

23. Описательная статистика в современных научных публикациях по географии. 24. Корреляционный анализ – возможности и перспективы использования в географических исследованиях. 25. Корреляционный анализ – возможности и перспективы использования в экономических исследованиях. 26. Обработка географических данных с помощью корреляционного анализа. 27. Обработка экономических данных с помощью корреляционного анализа. 28. Регрессионный анализ – возможности и перспективы использования в географических исследованиях.
Промежуточная аттестация
ОПК-9. Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности
Знает и понимает: современные программные средства и цифровые ресурсы для решения исследовательских задач Умеет: применять современные программные средства и цифровые ресурсы для решения исследовательских задач (стандартные статистические пакеты для осуществления обработки полученных эмпирических и экспериментальных данных) Владеет (навыками и/или опытом деятельности): навыками применения современных программных средств и цифровых ресурсов для решения исследовательских задач (стандартных статистических пакетов для осуществления обработки полученных эмпирических и экспериментальных данных)
Задания, необходимые для оценивания сформированности компетенции
Перечень вопросов для зачета
14. Проверка значимости коэффициента корреляции. 15. Значимость коэффициента детерминации. Проверка гипотезы H_0 о том, что все коэффициенты регрессии равны нулю. 16. Значимость оценок параметров регрессии. Проверка простой гипотезы H_0 о том, что оценка параметра регрессии равна истинному значению этого параметра регрессии. 17. Значимость оценок параметров регрессии. Проверка простой гипотезы H_0 о том, что оценки совокупности параметров регрессии не влияют на линию регрессии. 18. Значимость оценок параметров регрессии. Проверка простой гипотезы H_0 о том, что оценки двух параметров регрессии равны. 19. Значимость оценок параметров регрессии. Проверка простой гипотезы H_0 о том, что оценки параметров регрессии двух исследуемых процессов равны. 20. Цифровой спектральный анализ. Преобразование Фурье. 21. Коррелограммный метод оценки спектральной плотности мощности случайного процесса. 22. Периодограммный метод оценки спектральной плотности мощности случайного процесса. 23. Метод Уелча оценки спектральной плотности мощности случайного процесса. 24. Авторегрессионная (АР) модель последовательности наблюдений. Соотношение между параметрами АР-модели и корреляционной последовательностью. Оценивание коэффициентов АР-модели по методу наименьших квадратов (Уравнения Юла-Уолкера).

Лабораторная работа №1.

Тема: «Первичная обработка данных».

При проведении экспериментов фиксировались значения случайной величины X , характеризующей время простоя оборудования в ожидании ремонта (в часах).

Задание: произвести первичную обработку полученных опытных данных с целью изучения свойств случайной величины X .

1) Составим расчетную таблицу, в которой запишем вариационный ряд (элементы выборки в порядке возрастания признака) и произведем расчеты, необходимые для вычисления числовых характеристик.

Таблица 1 – Расчетная таблица

Номер п/п	Выборка, час.	Вариацион- ный ряд, x_i , час.	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$	$(x_i - \bar{x})^4$
1	6,72	0,21	-5,068	25,684	-130,163	659,6524
2	8,2	0,4	-4,885	23,864	-116,58	569,5061
3	0,4	0,64	-4,642	21,55	-100,036	464,3845
4	12,9	0,69	-4,595	21,118	-97,0476	445,9774
5	3,15	0,77	-4,509	20,33	-91,6644	413,3022
6	34,5	0,93	-4,346	18,892	-82,1113	356,8922
7	4,71	1,14	-4,142	17,158	-71,0714	294,3925
8	1,14	1,51	-3,772	14,226	-53,6592	202,3917
9	2,87	1,73	-3,55	12,601	-44,7323	158,7918
10	3,07	1,86	-3,422	11,709	-40,0658	137,0981
11	5,86	2	-3,277	10,741	-35,2003	115,3619
12	11,4	2,1	-3,179	10,105	-32,1227	102,1135
13	3,12	2,32	-2,964	8,7873	-26,0485	77,21641
14	0,21	2,32	-2,961	8,7695	-25,9695	76,9043
15	1,51	2,4	-2,885	8,3222	-24,0079	69,25838
16	2,76	2,76	-2,521	6,3562	-16,025	40,40161
17	0,93	2,87	-2,414	5,8281	-14,0699	33,96675
18	2,4	2,87	-2,409	5,8055	-13,988	33,70356
19	3,5	2,99	-2,294	5,2642	-12,0781	27,71187

20	5,29	3,07	-2,214	4,9038	-10,8593	24,0474
21	1,86	3,12	-2,163	4,6803	-10,1253	21,90503
22	4,99	3,15	-2,132	4,5437	-9,68547	20,6456
23	8,77	3,5	-1,776	3,1546	-5,60301	9,951645
24	1,73	3,6	-1,683	2,8313	-4,76407	8,016242
25	0,77	4,59	-0,691	0,477	-0,32948	0,227561
26	5,99	4,61	-0,669	0,4478	-0,29968	0,200545
27	7,95	4,71	-0,568	0,323	-0,18359	0,104347
28	2,87	4,99	-0,292	0,0854	-0,02497	0,007299
29	0,64	5,29	0,006	0,000036	0,0000002	1,66 ·10 ⁹
30	5,74	5,74	0,459	0,2103	0,096434	0,044222
31	0,69	5,86	0,582	0,3384	0,196888	0,11454
32	2,99	5,99	0,707	0,5003	0,35391	0,250336
33	4,59	6,72	1,439	2,0704	2,978986	4,286387
34	2,32	7,95	2,671	7,1361	19,06293	50,92364
35	2,32	8,2	2,924	8,55	25,00068	73,10309
36	2	8,77	3,492	12,193	42,57744	148,6755
37	2,1	11,4	6,097	37,173	226,6467	1381,866
38	4,61	12,9	7,608	57,874	440,2776	3349,414
39	30,1	30,1	24,78	614,27	15224,23	377323,6
40	3,6	34,5	29,26	855,97	25043,19	732688,6
Итого	211	211	0	1874,8	39956,09	1119385

2) Найдем размах выборки $R = x_{\max} - x_{\min} = 34,5 - 0,21 = 34,29$.

3) Длина интервала $h = \frac{R}{1 + 3,322 \lg n} = \frac{34,29}{1 + 3,322 \lg 40} = \frac{34,29}{1 + 3,322 \cdot 1,602} = 5,424$.

4) границы интервалов: $C_0 = x_{\min} = 0,21$,

$C_1 = C_0 + h = 0,21 + 5,424 = 5,634$, $C_2 = C_1 + h = 5,634 + 5,424 = 11,058$,

$C_3 = C_2 + h = 11,058 + 5,424 = 16,482$, $C_4 = C_3 + h = 16,482 + 5,424 = 21,906$, $C_5 = C_4 + h = 21,906 + 5,424 = 27,33$,

$C_6 = C_5 + h = 27,33 + 5,424 = 32,754$, $C_7 = C_6 + h = 32,754 + 5,424 = 38,178 > x_{\max}$.

5) Построим интервальный статистический ряд:

Таблица 2 – Интервальный статистический ряд

Границы интервалов $[C_i, C_{i+1})$, час.	Частоты m_i	Частоты m_i / n	Накопленные частоты
[0,21; 5,634)	29	29/40	29/40
[5,634; 11,058)	7	7/40	36/40
[11,058; 16,482)	2	2/40	38/40
[16,482; 21,906)	0	0/40	38/40
[21,906; 27,33)	0	0/40	38/40
[27,33; 32,754)	1	1/40	39/40
[32,754; 38,178)	1	1/40	1
Итого	40	1	

6) Вычислим числовые характеристики.

В качестве оценки **математического ожидания** используется среднее арифметическое $\bar{x}$ наблюдаемых значений. Эта статистика называется **выборочным средним**.

$$\hat{M}[X] = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{211}{40} = 5,281 (\text{часов})$$

Для оценивания по выборочным данным моды распределения, используется то значение сгруппированного статистического ряда $\hat{x}_{\text{mod}}$, которому соответствует наибольшее значение частоты. По интервальному статистическому ряду определяется модальный интервал, в который попало **наибольшее число элементов выборки**, и в качестве точечной оценки моды может использоваться среднее значение этого интервала.

$$\hat{x}_{\text{mod}} = \frac{0,21 + 5,634}{2} = 2,922(\text{часа})$$

Для определения выборочного значения медианы используется вариационный ряд. В качестве оценки медианы $\hat{x}_{\text{med}}$ принимают средний (т. е. $\frac{1}{2}(n+1)$ -й) член этого ряда, если значение n – нечётно и среднее арифметическое между двумя средними (т. е. между $\frac{1}{2}n$ -м и $\left(\frac{1}{2}n+1\right)$ -м) членами этого ряда, если n – чётно. В нашем случае объем выборки $n = 40$ – четное, т.е. в качестве оценки медианы примем

$$= \frac{x_{20} + x_{21}}{2} = \frac{3,07 + 3,12}{2} = 3,095(\text{часа})$$

В качестве оценки дисперсии используется

$$\hat{D}[X] = \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1874,8}{40-1} = 48,072(\text{часа}^2)$$

Оценка среднего квадратического отклонения

$$\hat{\sigma}[X] = \sqrt{\hat{D}[X]} = \sqrt{48,072} = 6,933(\text{часа})$$

$$\hat{V} = \frac{\hat{\sigma}}{\bar{x}} \cdot 100\% = \frac{6,933}{5,281} \cdot 100 = 131,3\%$$

Оценка коэффициента вариации

Оценка коэффициента асимметрии

$$\hat{A}[X] = \hat{\beta}_1 = \frac{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^3}{\hat{\sigma}^3} = \frac{39956,09}{39 \cdot 6,933^3} = 3,074$$

Оценка коэффициента эксцесса

$$\hat{E}[X] = \hat{\beta}_2 = \frac{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^4}{\hat{\sigma}^4} - 3 = \frac{1119385}{39 \cdot 6,933^4} - 3 = 9,423$$

7) Для приближённого построения эмпирической функции распределения воспользуемся соотношением:

$$\hat{F}[x] = \sum_{x_i < x} \frac{m_i}{n}$$

8) Построим гистограмму частот и эмпирическую функцию распределения.

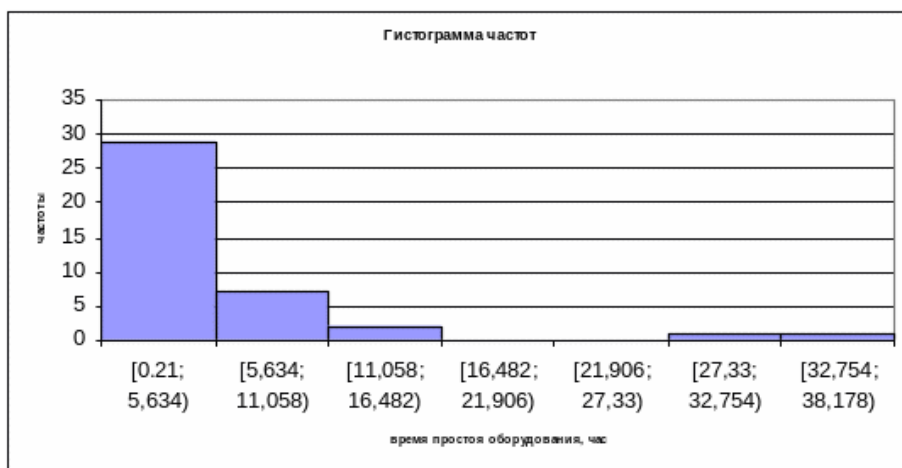


Рисунок 1 – Гистограмма частот

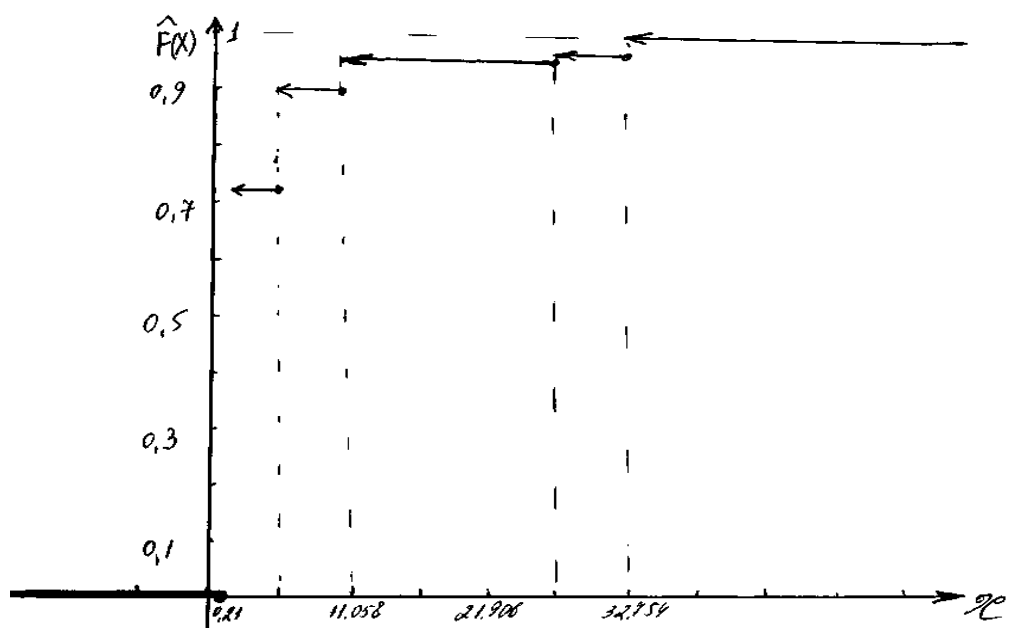


Рисунок 2 – Функция распределения

Вывод. В результате исследования выборки значений непрерывной случайной величины, характеризующей время простоя оборудования в ожидании ремонта, получили следующие результаты, час: минимальное время простоя – 0,21, максимальное – 34,54, среднее значение времени простоя оборудования – 5,28, наиболее вероятное время простоя оборудования – 2,922, средневероятное – 3,095, среднеквадратическое отклонение времени простоя оборудования от среднего значения составило 6,933. Оценка коэффициента вариации составила 131,3%, что указывает на большую колеблемость признака относительно среднего значения, оценка коэффициента асимметрии составила 3,074, оценка коэффициента эксцесса составила 9,423.

Лабораторная работа №2.

Тема: «Подбор закона распределения одномерной случайной величины»

Цель работы: изучить методику применения критерия χ^2 Пирсона для проверки гипотезы о виде закона распределения случайной величины.

Задание: с помощью критерия проверить согласование выдвинутой гипотезы о виде закона распределения исследуемой случайной величины с имеющимися выборочными данными.

Алгоритм применения критерия χ^2 для проверки гипотезы о виде закона распределения исследуемой случайной величины.

1 Выборочные данные представляются в виде интервального или сгруппированного статистического ряда.

2 Выбирается уровень значимости α .

3 Формулируется гипотеза о виде закона распределения исследуемой случайной величины.

4 Вычисляются вероятности p_i попадания значений случайной величины X в рассматриваемые разряды разбиения: $p_i = P(C_i \leq X < C_{i+1}) = F(C_{i+1}) - F(C_i)$, ($i = 1, 2, \dots, k$), где $F(x)$ – гипотетическая функция распределения случайной величины X .

Замечание. Если изучается непрерывная случайная величина, то при вычислении значений p_i **необходимо изменить границы** первого и последнего частичных интервалов разбиения таким образом, чтобы учесть все возможные значения, которые может принять случайная величина предполагаемого класса. В зависимости от конкретного вида проверяемой гипотезы границы частичных интервалов необходимо изменить следующим образом:

Вид закона распределения	Первый интервал разбиения	Последний интервал разбиения
Равномерный	$[\hat{a}, C_2)$	$[C_k, \hat{b})$
Экспоненциальный	$[0, C_2)$	$[C_k, \infty)$
Нормальный	$(-\infty, C_2)$	

5 Определяются значения теоретических частот np_i ($i = 1, 2, \dots, k$). При необходимости для обеспечения условия $np_i \geq 3$ (если объем выборки $n < 50$), $np_i \geq 5$ (если объем выборки $n \geq 50$), объединяются несколько соседних разрядов разбиения.

$$\chi^2 = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i}$$

6 Вычисляется наблюдаемое значение критерия χ^2 :

7 По таблицам квантилей распределения χ^2 определяется критическое значение $\chi_{\alpha, v}^2$, соответствующее заданному уровню значимости α и числу степеней свободы $v = k - r - 1$.

Если расчётное значение критерия попадает в критическую область, т. е. $\chi^2 > \chi_{\alpha, v}^2$, то проверяемая гипотеза отвергается (при этом вероятность отклонения верной гипотезы равна α).

В случаях, когда наблюдаемое значение χ^2 не превышает критического $\chi^2 \leq \chi_{\alpha, v}^2$, считают, что выдвинутая гипотеза не противоречит опытным данным. Подчеркнем, что

полученный результат свидетельствует лишь о приемлемом согласовании проверяемой гипотезы с имеющимися выборочными данными и, в общем случае, не является доказательством истинности этой гипотезы.

Пример 2.1. По таблице, полученной в лабораторной работе №1 и по гистограмме частот выдвигаем нулевую гипотезу о виде закона распределения случайной величины X (времени простоя оборудования в ожидании ремонта).

H_0 : Случайная величина (время простоя оборудования в ожидании ремонта) распределена по показательному (экспоненциальному) закону.

Выбираем уровень значимости $\alpha = 0,05$.

Вычислим вероятности p_i попадания значений случайной величины X в рассматриваемые разряды разбиения по формуле: $p_i = P(C_i \leq X < C_{i+1}) = F(C_{i+1}) - F(C_i) = (1 - e^{-\lambda C_{i+1}}) - (1 - e^{-\lambda C_i}) = e^{-\lambda C_i} - e^{-\lambda C_{i+1}}$.

Проверим гипотезу H_0 с помощью критерия согласия Хи-квадрат Пирсона.

Вычислим параметр $\lambda = \frac{1}{\bar{x}} = \frac{1}{5,281} = 0,189358 = 0,189$.

Так как изучается непрерывная случайная величина, то при вычислении значений **необходимо изменить границы** первого и последнего частичных интервалов разбиения. В нашем случае проверяется гипотеза о показательном законе распределения.

Вид закона распределения	Первый интервал разбиения	Последний интервал разбиения
Экспоненциальный		

Вычислим вероятности по

формуле $p_i = P(C_i \leq X < C_{i+1}) = e^{-\lambda C_i} - e^{-\lambda C_{i+1}}$ $p_i = P(C_i \leq X < C_{i+1}) = e^{-0,189 C_i} - e^{-0,189 C_{i+1}}$.

Пример расчета:

$$p(0 \leq X < 5,634) = e^{-0,189 \cdot 0} - e^{-0,189 \cdot 5,634} = 1 - 0,344788 = 0,655212 = 0,655.$$

Для того, чтобы облегчить расчеты, можно с помощью пакета

программ **XL** выполнить промежуточные расчеты, которые необходимо оформить в виде таблицы:

Таблица 1 - Расчетная таблица вероятностей

Граница интервала C_i	$-\lambda \cdot C_i$	$e^{-\lambda C_i} = e^{-0,189 C_i}$	$p_i = e^{-\lambda C_i} - e^{-\lambda C_{i+1}}$	np_i
0	0	1	0,655212	0,655212
5,634	-1,06483	0,344788	0,221096	0,221096

11,058	- 2,08996	0,123692	0,079318	0,079318
16,482	-3,1151	0,044374	0,028455	0,028455
21,906	- 4,14023	0,015919	0,010208	0,010208
27,33	- 5,16537	0,005711	0,003662	0,003662
32,754	- 6,19051	0,002049	0,002049	0,002049
∞	-	0	-	-
Итого	-	-	1	40

Таблица 2 – Расчет χ^2

Границы интервала в	Частоты эмпирические	Вероятность и p_i	Частоты теоретические $np_i = 40 \cdot p_i$	$\frac{(m_i - np_i)^2}{np_i}$
[0; 5,634)	29	0,655	26,21	$\frac{(29 - 26,21)^2}{26,21} = 0,297$
[5,634; 11,058)	7	0,221	8,844	$\frac{(7 - 8,844)^2}{8,844} = 0,385$
[11,058; 16,482)	2	0,079	3,173	$\frac{(4 - 4,947)^2}{4,947} = 0,181$
[16,482; 21,906)	0	0,028	1,138	
[21,906; 27,33)	0	0,01	0,408	
[27,33; 32,754)	1	0,004	0,146	
[32,754;)	1	0,002	0,082	$0,863 = \chi^2$
Итого	40	1	40	

Вычислим число степеней свободы $\chi = k - r - 1 = 3 - 1 - 1 = 1$, где $k = 3$ – число интервалов в таблице 2 после объединения, $r = 1$ – число параметров выбранного закона распределения – в нашем случае показательный закон (один параметр).

По таблицам квантилей распределения χ^2 определяется критическое значение $= \chi_{0,05,1}^2 = 3,841$, соответствующее заданному уровню значимости $\alpha = 0,05$ и числу степеней свободы $\chi = 1$.

Вывод. Сравниваем полученное значение в таблице $\chi^2 = 0,863$ с табличным $\chi^2_{\alpha, \nu} = 3,841$. Так как расчетное = 0,863 меньше, чем табличное = 3,841, то гипотеза о показательном законе распределения подтвердилась.

Пример.2.2

При проведении экспериментов фиксировались значения случайной величины X , характеризующей цены на зимнюю обувь (в у.е.).

Задание: произвести первичную обработку полученных опытных данных с целью изучения свойств случайной величины X , построить гистограмму частот. По гистограмме частот выдвинуть нулевую гипотезу о виде закона распределения случайной величины и проверить ее с помощью критерия согласия Пирсона.

- 1) Составим расчетную таблицу, в которой запишем вариационный ряд (элементы выборки в порядке возрастания признака) и произведем расчеты, необходимые для вычисления числовых характеристик.

Таблица 1- Расчетная таблица

Номер п/п	Выборка, у.е.	Вариационный ряд, у.е.,		
1	126	85	- 42,16	1777,466
2	105	89	- 38,16	1456,186
3	134	89	- 38,16	1456,186
4	134	97	- 30,16	909,6256
5	152	100	- 27,16	737,6656
6	97	102	- 25,16	633,0256
7	128	104	- 23,16	536,3856
8	128	104	- 23,16	536,3856
9	153	105	- 22,16	491,0656
10	128	105	- 22,16	491,0656
11	116	106	- 21,16	447,7456

12	89	109	- 18,16	329,7856
13	105	113	- 14,16	200,5056
14	121	114	- 13,16	173,1856
15	104	116	- 11,16	124,5456
16	120	118	-9,16	83,9056
17	155	118	-9,16	83,9056
18	183	119	-8,16	66,5856
19	150	120	-7,16	51,2656
20	154	121	-6,16	37,9456
21	140	121	-6,16	37,9456
22	140	123	-4,16	17,3056
23	85	123	-4,16	17,3056
24	137	126	-1,16	1,3456
25	104	128	0,84	0,7056
26	134	128	0,84	0,7056
27	146	128	0,84	0,7056
28	118	132	4,84	23,4256
29	106	134	6,84	46,7856
30	141	134	6,84	46,7856
31	136	134	6,84	46,7856
32	141	134	6,84	46,7856
33	100	136	8,84	78,1456
34	121	137	9,84	96,8256
35	118	139	11,84	140,1856
36	89	140	12,84	164,8656

37	123	140	12,84	164,8656
38	139	141	13,84	191,5456
39	132	141	13,84	191,5456
40	158	143	15,84	250,9056
41	114	146	18,84	354,9456
42	151	150	22,84	521,6656
43	123	151	23,84	568,3456
44	119	152	24,84	617,0256
45	143	153	25,84	667,7056
46	134	154	26,84	720,3856
47	160	155	27,84	775,0656
48	109	158	30,84	951,1056
19	113	160	32,84	1078,466
50	102	183	55,84	3118,106
Итого	6358	6358	0	21562,72

2) Вычислим числовые характеристики.

$$\hat{M}[X] = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{6358}{50} = 127,16 (y.e.)$$

Выборочное среднее:

Мода: .

$$\hat{x}_{\text{mod}} = \frac{127 + 141}{2} = 134 (y.e.)$$

$$\text{Медиана: } = \frac{x_{25} + x_{26}}{2} = \frac{128 + 128}{2} = 128 (y.e.)$$

В качестве оценки дисперсии используется статистика

$$= \frac{21562,72}{50 - 1} = 440,056 (y.e.^2)$$

Оценка среднего квадратического отклонения

$$= \sqrt{440,056} = 20,978 (y.e.)$$

Оценка коэффициента вариации

$$\hat{V} = \frac{\hat{\sigma}}{\bar{x}} \cdot 100\% = \frac{20,978}{127,16} \cdot 100 = 16,5\%$$

Найдем размах выборки $R = X_{\max} - X_{\min} = 183 - 85 = 98$.

3) Вычислим длину интервала $= \frac{98}{1 + 3,322 \lg 50} = 14$.

4) Границы интервалов:

$$C_0 = X_{\min} = 85, = 85 + 14 = 99,$$

$$= 99 + 14 = 113, = 113 + 14 = 127,$$

$$= 127 + 14 = 141, = 141 + 14 = 155, = 155 + 14 = 169, = 169 + 14 = 183 = X_{\max}.$$

Построим гистограмму частот.



Рисунок 1 – Гистограмма частот

H_0 : Случайная величина (цены на товары (в у.е.)) распределена по нормальному закону.

Выбираем уровень значимости.

Так как изучается непрерывная случайная величина, то при вычислении значений **необходимо изменить границы** первого и последнего частичных интервалов разбиения. В нашем случае проверяется гипотеза о нормальном законе распределения.

Вид закона распределения	Первый интервал разбиения	Последний интервал разбиения
Нормальный		

Вычислим вероятности p_i попадания значений случайной величины X в

рассматриваемые разряды разбиения по формуле:

$$\Phi\left(\frac{C_{i+1} - \bar{X}}{\hat{\sigma}[X]}\right) - \Phi\left(\frac{C_i - \bar{X}}{\hat{\sigma}[X]}\right).$$

Проверим гипотезу с помощью критерия согласия Хи-квадрат Пирсона.

Вычислим параметр $\bar{X} = 127,16$, $\hat{\sigma}[X] = 20,978$.

Вычислим вероятности по

$$\text{формуле } p_i = P(C_i \leq X < C_{i+1}) = \Phi\left(\frac{C_{i+1} - 127,16}{20,978}\right) - \Phi\left(\frac{C_i - 127,16}{20,978}\right)$$

$$p(-\infty \leq X < 99) = \Phi\left(\frac{99 - 127,16}{20,978}\right) - \Phi\left(\frac{-\infty - 127,16}{20,978}\right) = \Phi(-1,342) - \Phi(-\infty) = 0,4099 + 0,5 = 0,0901,$$

$$p(99 \leq X < 113) = \Phi\left(\frac{113 - 127,16}{20,978}\right) - \Phi\left(\frac{99 - 127,16}{20,978}\right) = \Phi(-0,675) - \Phi(-1,342) = -0,2517 + 0,4099 = 0,1582,$$

$$p(113 \leq X < 127) = \Phi\left(\frac{127 - 127,16}{20,978}\right) - \Phi\left(\frac{113 - 127,16}{20,978}\right) = \Phi(-0,008) - \Phi(-0,675) = -0,0040 + 0,2517 = 0,2477,$$

$$p(127 \leq X < 141) = \Phi\left(\frac{141 - 127,16}{20,978}\right) - \Phi\left(\frac{127 - 127,16}{20,978}\right) = \Phi(0,660) - \Phi(-0,008) = 0,2454 + 0,0040 = 0,2494,$$

$$p(141 \leq X < 155) = \Phi\left(\frac{155 - 127,16}{20,978}\right) - \Phi\left(\frac{141 - 127,16}{20,978}\right) = \Phi(1,327) - \Phi(0,660) = 0,4082 - 0,2454 = 0,1628,$$

$$p(155 \leq X < 169) = \Phi\left(\frac{169 - 127,16}{20,978}\right) - \Phi\left(\frac{155 - 127,16}{20,978}\right) = \Phi(1,995) - \Phi(1,327) = 0,4772 - 0,4082 = 0,069,$$

$$p(169 \leq X < +\infty) = \Phi\left(\frac{+\infty - 127,16}{20,978}\right) - \Phi\left(\frac{169 - 127,16}{20,978}\right) = \Phi(+\infty) - \Phi(1,995) = 0,5 - 0,4772 = 0,0228.$$

Таблица 2 – Расчет Хи-квадрат

Границы интервалов	Частоты эмпирические	Вероятности	Частоты теоретические $np_i = 50 \cdot p_i$	
$(-\infty, 99]$	4	0,0901	4,505	0,028
$(99, 113]$	9	0,1582	7,91	
$(113, 127]$	11	0,2477	12,385	0,155
$(127, 141]$	16	0,2494	12,47	0,999
$(141, 155]$	7	0,1628	12,73	0,586
$(155, 169]$	2	0,069		
$(169, +\infty)$	1	0,0228		
итого	50	1	50	1,768=

Вывод. По таблицам квантилей распределения χ^2 определяется критическое значение $= 3,841$, соответствующее заданному уровню значимости $\alpha=0,05$ и числу степеней свободы $\nu=4-2-1=1$. Сравниваем полученное значение $= 1,768$ с табличным значением $= 3,841$. Так как расчетное $= 1,768$ меньше, чем табличное $= 3,841$, то гипотеза о нормальном законе распределения подтвердилась.

Лабораторная работа №3.

Тема: «Построение интервальных оценок параметров распределения»

Цель работы: изучить методику построения интервальных оценок

параметров распределения вероятностей случайной величины.

Задание: построить интервальные оценки математического ожидания и дисперсии исследуемой случайной величины.

Пример 1 При проведении экспериментов фиксировались значения случайной величины X , характеризующей цены на зимнюю обувь (в у.е.). В результате вычислений, проведенных в Лабораторной работе № 1, были получены следующие числовые характеристики:

выборочное среднее: .

Оценка среднего квадратического отклонения

= .

Объем выборки составил $n = 50$. Требуется с доверительной вероятностью $P = 0,95$ определить интервальную оценку:

а) для средней цены на зимнюю обувь;

б) для дисперсии цены на зимнюю обувь;

в) для среднего квадратического отклонения цены на зимнюю обувь.

Решение. а) Средняя цена на обувь характеризуется генеральной средней a . Требуется найти интервальную оценку $(\hat{a}_1; \hat{a}_2)$ параметра a с доверительной вероятностью .

Применяем формулу $\left\{ \bar{x} - t_{(\alpha/2; n-1)} \frac{\hat{\sigma}}{\sqrt{n-1}} < a < \bar{x} + t_{(\alpha/2; n-1)} \frac{\hat{\sigma}}{\sqrt{n-1}} \right\}$, где

$\bar{x} = 127,16$ (у.е.) , $\hat{\sigma}[X] = 20,978$ (у.е.) , $n = 50$, $\alpha = 0,05$,

значение $t_{\alpha/2; n-1}$ определяем по таблицам распределения Стьюдента для $\alpha/2 = 0,025$ и $n-1 = 50-1 = 49$. $t(0,025; 49) = 2,009$. Подставим найденные значения в формулу:

$$\left\{ 127,16 - 2,009 \frac{20,978}{\sqrt{49}} < a < 127,16 + 2,009 \frac{20,978}{\sqrt{49}} \right\}$$

$$[121,139 < a < 133,181] \text{ у.е.}$$

Таким образом, с вероятностью можно гарантировать, что средняя цена зимней обуви находится в пределах: $[121,139 \text{ у.е.} < a < 133,181 \text{ у.е.}]$.

б) определим интервальную оценку для дисперсии цены на зимнюю обувь.

Интервальная оценка дисперсии

$$\frac{(n-1)\sigma^2}{\chi^2_{(\alpha/2; n-1)}} < \sigma^2 < \frac{(n-1)\sigma^2}{\chi^2_{(1-\alpha/2; n-1)}}.$$

По таблице процентных точек χ^2 -распределения (см. приложение Г) найдем

$$\chi^2_{(\alpha/2; n-1)} = \chi^2_{(0,025; 49)} = 59,342;$$

$$\chi^2_{(1-\alpha/2; n-1)} = \chi^2_{(0,975; 49)} = 24,433.$$

Следовательно,

$$\sigma^2 \in \left(\frac{49 \cdot 440,077}{59,342}; \frac{49 \cdot 440,077}{24,433} \right).$$

Значит с доверительной вероятностью можно утверждать, что истинное значение дисперсии будет находиться в интервале

$$363,381 < \sigma^2 < 882,567.$$

в) С доверительной вероятностью можно утверждать, что истинное значение среднего квадратического отклонения σ будет находиться в интервале

$$19,063 \text{ у.е.} < \sigma < 29,708 \text{ у.е.}$$

Порядок выполнения работы

1 В лабораторной работе № 1 вы выполнили обработку статистических данных и вычислили точечную оценку $\bar{X}$ среднего значения исследуемой случайной величины X и точечную оценку $\hat{\sigma}$ среднего квадратического отклонения X . Перепишите из лабораторной работы № 1 значения величин n , и σ .

2 Постройте вручную интервальные оценки для неизвестных истинных значений $M[X]$ и $\sigma[X]$.

3 Вычислите интервальные оценки для σ на ЭВМ..

4 Сделайте вывод.

Лабораторная работа № 4 Построение регрессионной модели системы двух случайных величин

Цель работы: изучить основные методы регрессионного и корреляционного анализа; исследовать зависимость между двумя случайными величинами, заданными выборками.

Задание: по виду корреляционного поля сделать предположение о форме регрессионной зависимости между двумя случайными величинами; используя метод наименьших квадратов, найти параметры уравнения регрессии; оценить качество описания зависимости полученным уравнением регрессии.

Пример . По результатам пятнадцати совместных измерений веса грузового поезда, t , и соответствующего времени нахождения поезда на участке Y , τ , представленных в таблице 1, следует исследовать зависимость между данными величинами. Необходимо определить коэффициенты уравнения регрессии методом наименьших квадратов, оценить тесноту связи между величинами, проверить значимость коэффициента корреляции и спрогнозировать время нахождения поезда на участке при заданном весе поезда (5200 т).

Решение. На величину времени нахождения поезда на участке Y , помимо веса X , влияние оказывает профиль и качество железнодорожного полотна, качество подвижного состава, направление и скорость ветра и другие факторы. Поэтому зависимость между величиной времени нахождения поезда на участке Y и веса поезда X является статистической: при одном весе поезда при различных дополнительных условиях время нахождения поезда на участке может принимать различные значения. Для определения вида регрессионной зависимости построим корреляционное поле.

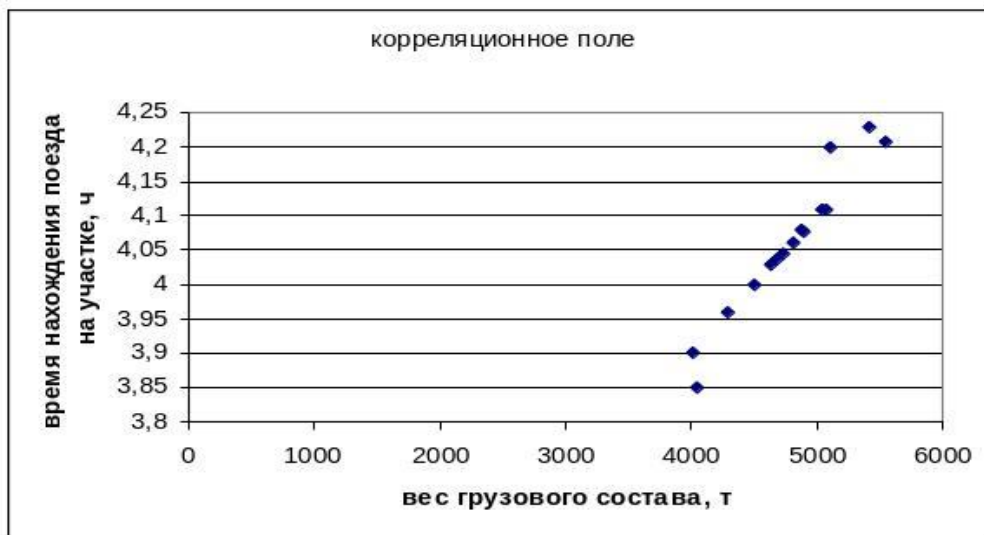


Рис.1. Корреляционное поле

Характер расположения точек на диаграмме рассеяния позволяет сделать предположение о линейной регрессионной зависимости $\hat{y}_i = \beta_0 + \beta_1 \cdot x_i$.

Таблица 1 - Результаты промежуточных вычислений

Вес грузового состава, т, x_i	Время нахождения поезда на участке, час., y_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x}) \cdot (y_i - \bar{y})$
5100,58	4,2	327,019	106941,2	0,14007	0,01962	45,8044
4885,41	4,078	111,849	12510,12	0,01807	0,00033	2,02073
5416,94	4,23	643,379	413936,1	0,17007	0,02892	109,417
4496,66	4,001	-276,901	76674,35	-0,0589	0,00347	16,3187
4722,08	4,044	-51,4813	2650,328	-0,0159	0,00025	0,82027
5537,91	4,208	764,349	584228,9	0,14807	0,02192	113,175
5074,01	4,11	300,449	90269,4	0,05007	0,00251	15,0425

4807,09	4,062	33,5287	1124,171	0,00207	4,3E-06	0,06929
4046,02	3,85	- 727,541	529316,4	-0,2099	0,04407	152,735
4683,93	4,037	- 89,6313	8033,776	-0,0229	0,00053	2,05555
4872,42	4,08	98,8587	9773,036	0,02007	0,0004	1,98376
4003,22	3,9	- 770,341	593425,8	-0,1599	0,02558	123,203
4628,01	4,03	- 145,551	21185,19	-0,0299	0,0009	4,35684
4293,44	3,96	- 480,121	230516,5	-0,0999	0,00999	47,9801
5035,7	4,109	262,139	68716,68	0,04907	0,00241	12,8623
Итого 71603,42	60,899	0	2749302	0	0,1609	647,845

Найдем уравнение прямой линии методом наименьших квадратов .

Средний вес грузового состава:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{71603,42}{15} = 4773,561 \text{ м}$$

Среднее значение времени нахождения поезда на участке:

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{60,899}{15} = 4,060 \text{ ч}$$

Коэффициенты уравнения:

$$\beta_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{647,845}{2749302} = 0,00024$$

$$\beta_0 = \bar{y} - \bar{x} \cdot \beta_1 = 4,060 - 4773,561 \cdot 0,00024 = 2,935$$

Уравнение регрессии имеет вид : $\hat{y}_i = 2,935 + 0,00024 \cdot x_i$.

Для линейной связи коэффициенты:

β_0 - постоянная регрессии, показывает точку пересечения прямой с осью ординат

β_1 - коэффициент регрессии, показывает меру зависимости переменных y от x , указывает среднюю величину изменения переменной y при изменении x на одну единицу, знак β_1 определяет направление этого изменения .

Вычислим линейный коэффициент корреляции

$$\hat{r} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{647,845}{\sqrt{2749302} \sqrt{0,1609}} = 0,974$$

Таблица 2 - Расчет значений времени нахождения поезда на участке по уравнению регрессии

Вес грузового состава, т,	Время нахождения поезда на участке, час.,	
5100,58	4,2	4,137
4885,41	4,078	4,0863
5416,94	4,23	4,2115
4496,66	4,001	3,9947
4722,08	4,044	4,0478
5537,91	4,208	4,24
5074,01	4,11	4,1307
4807,09	4,062	4,0678
4046,02	3,85	3,8885
4683,93	4,037	4,0388
4872,42	4,08	4,0832
4003,22	3,9	3,8784
4628,01	4,03	4,0256
4293,44	3,96	3,9468
5035,7	4,109	4,1217
Итого 71603,42	60,899	60,899

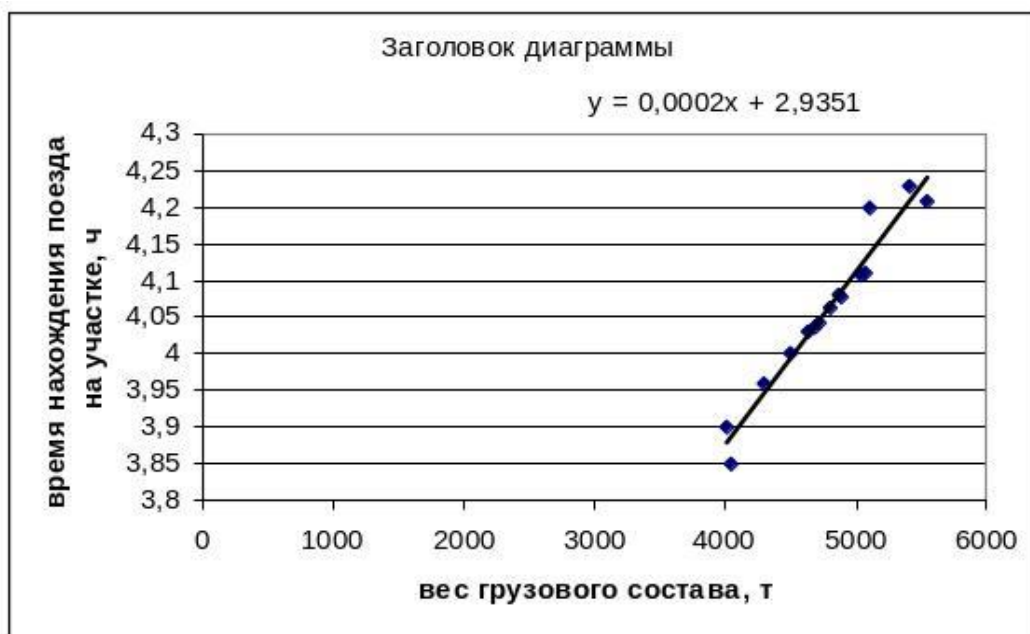


Рис.2. Корреляционное поле и линия регрессии

Спрогнозируем время нахождения поезда на участке при заданном весе грузового состава (5200 т).

$$\hat{y}_i = 2,935 + 0,00024 \cdot 5200 = 4,183 \text{ ч.}$$

Качественная оценка тесноты связи между величинами выявляется по шкале Чеддока (таблица 3).

Таблица 3 - Шкала Чеддока

Теснота связи	Значение коэффициента корреляции при наличии	
	прямой связи	обратной связи
Слабая	0,1–0,3	(-0,1)–(-0,3)
Умеренная	0,3–0,5	(-0,3)–(-0,5)
Заметная	0,5–0,7	(-0,5)–(-0,7)
Высокая	0,7–0,9	(-0,7)–(-0,9)
Весьма высокая	0,9–0,99	(-0,9)–(-0,99)

Вывод. Линейный коэффициент корреляции характеризует тесноту связи между двумя коррелируемыми признаками в случае наличия между ними линейной зависимости.

Т.к. $r_{xy} = 0,974$, то можно говорить о том, что между величинами X и Y существует линейная прямая, весьма высокая связь.

Чтобы сделать статистический вывод о значимости коэффициента корреляции (при проверке линейности регрессионной зависимости) выдвигается нулевая гипотеза об

отсутствии линейной зависимости между исследуемыми с. в. против альтернативной гипотезы о наличии линейной связи.

$$H_0: r_{xy} = 0,$$

$$H_a: r_{xy} \neq 0.$$

Если гипотеза H_0 отклоняется, то считается, что уравнение регрессии Y по X действительно имеет линейный вид.

Для проверки гипотезы H_0 вычисляется t -статистика

$$t = \hat{r} \sqrt{\frac{n-2}{1-\hat{r}^2}} = 0,974 \sqrt{\frac{15-2}{1-0,974^2}} = 15,501.$$

При условии справедливости гипотезы H_0 рассчитанная t -статистика имеет распределение Стьюдента с $n-2$ степенями свободы. Найденное значение $t = 6,247$ сравнивается с критическим значением $t_{\alpha, n-2}$ при $\alpha = n-2 = 15-2 = 13$ степенях свободы (приложение Д). В нашем случае $t_{\alpha, n-2} = t_{0.05, 13} = 1,771$. Так как расчетное значение 15,501 по абсолютной величине превосходит табличное 1,771 для заданного уровня значимости, то нулевая гипотеза H_0 о линейной независимости двух с. в. отклоняется.

Характер расположения точек на диаграмме рассеяния позволяет сделать предположение о параболической регрессионной зависимости

$$\bar{y}(x) = \beta_0 + \beta_1 x + \beta_2 x^2.$$

Оценки параметров β_0 , β_1 и β_2 найдем методом наименьших квадратов. Для этого составим функцию $S(\beta_0, \beta_1, \beta_2)$, которая в случае параболической регрессии примет вид

$$S(\beta_0, \beta_1, \beta_2) = \sum_{i=1}^n (y_i - \bar{y}(x_i))^2 = \sum_{i=1}^n [y_i - (\beta_0 + \beta_1 x_i + \beta_2 x_i^2)]^2.$$

Для отыскания оценок параметров β_0 , β_1 и β_2 , минимизирующих функцию $S(\beta_0, \beta_1, \beta_2)$, составим и решим систему нормальных уравнений:

$$S = \sum_{i=1}^n (y_i - \bar{y}(x_i))^2 = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2 = S(\beta_0, \beta_1, \beta_2)$$

$$\begin{cases} \frac{\partial S(\beta_0, \beta_1, \beta_2)}{\partial \beta_0} = 0; \\ \frac{\partial S(\beta_0, \beta_1, \beta_2)}{\partial \beta_1} = 0; \\ \frac{\partial S(\beta_0, \beta_1, \beta_2)}{\partial \beta_2} = 0 \end{cases}$$

$$\begin{cases} 2(-1) \sum_{i=1}^n [y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2] \cdot 1 = 0; \\ 2(-1) \sum_{i=1}^n [y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2] \cdot x_i = 0; \\ 2(-1) \sum_{i=1}^n [y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2] \cdot x_i^2 = 0; \end{cases}$$

Разделим обе части уравнений (1,2,3) на (-2)

$$\begin{cases} \sum_{i=1}^n [y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2] = 0; \\ \sum_{i=1}^n [y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2] x_i = 0; \\ \sum_{i=1}^n [y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2] x_i^2 = 0; \end{cases}$$

$$\begin{cases} \sum_{i=1}^n y_i - \beta_0 n - \beta_1 \sum_{i=1}^n x_i - \beta_2 \sum_{i=1}^n x_i^2 = 0; \\ \sum_{i=1}^n y_i x_i - \beta_0 \sum_{i=1}^n x_i - \beta_1 \sum_{i=1}^n x_i^2 - \beta_2 \sum_{i=1}^n x_i^3 = 0; \\ \sum_{i=1}^n y_i x_i^2 - \beta_0 \sum_{i=1}^n x_i^2 - \beta_1 \sum_{i=1}^n x_i^3 - \beta_2 \sum_{i=1}^n x_i^4 = 0 \end{cases}$$

$$\begin{cases} \beta_0 n + \beta_1 \sum_{i=1}^n x_i + \beta_2 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i; \\ \beta_0 \sum_{i=1}^n x_i + \beta_1 \sum_{i=1}^n x_i^2 + \beta_2 \sum_{i=1}^n x_i^3 = \sum_{i=1}^n y_i x_i; \\ \beta_0 \sum_{i=1}^n x_i^2 + \beta_1 \sum_{i=1}^n x_i^3 + \beta_2 \sum_{i=1}^n x_i^4 = \sum_{i=1}^n y_i x_i^2 \end{cases}$$

Для вычисления значений сумм, входящих в систему уравнений, составим расчетную таблицу 4.

Таблица 4 – Результаты промежуточных вычислений

Вес грузового состава, т,	Время нахождения поезда на участке, час.,	x_i^2	x_i^3	x_i^4	$x_i y_i$	$x_i^2 y_i$
5100,58	4,2	26015916	$1,32696 \cdot 10^{11}$	$6,76828 \cdot 10^{14}$	21422,44	109266848,6
4885,41	4,078	23867231	1,16601	5,69645	19922,7	97330567,48
5416,94	4,23	29343239	1,58951	8,61026	22913,66	124121900,8
4496,66	4,001	20219951	90922245563	4,08846	17991,14	80900024,57
4722,08	4,044	22298040	1,05293	4,97203	19096,09	90173271,84
5537,91	4,208	30668447	1,69839	9,40554	23303,53	129052825,7
5074,01	4,11	25745577	1,30633	6,62835	20854,18	105814323,4
4807,09	4,062	23108114	1,11083	5,33985	19526,4	93865160,16
4046,02	3,85	16370278	66234471548	2,67986	15577,18	63025569,69
4683,93	4,037	21939200	1,02762	4,81329	18909,03	88568551,39
4872,42	4,08	23740477	1,15674	5,6361	19879,47	96861144,76
4003,22	3,9	16025770	64154684454	2,56825	15612,56	62500504,44
4628,01	4,03	21418477	99124923705	4,58751	18650,88	86316460,54
4293,44	3,96	18433627	79143671651	3,39799	17002,02	72997163,05
5035,7	4,109	25358274	1,27697	6,43042	20691,69	104197149,9
Итого 71603,42	60,899	$3,45 \cdot 10^8$	$1,67081 \cdot 10^{12}$	$8,16226 \cdot 10^{15}$	291353	1404991466

После подстановки значений система уравнений примет вид:

$$\begin{cases} 15\beta_0 + 71603,42\beta_1 + 3,45 \cdot 10^8 \beta_2 = 60,899 \\ 71603,42\beta_0 + 3,45 \cdot 10^8 \beta_1 + 1,67081 \cdot 10^{12} \beta_2 = 291353 \\ 3,45 \cdot 10^8 \beta_0 + 1,67081 \cdot 10^{12} \beta_1 + 8,16226 \cdot 10^{15} \beta_2 = 1404991466 \end{cases}$$

Решив систему уравнений известными методами (методом Крамера, методом Гаусса, методом обратной матрицы) или с помощью MATHCAD, получим следующее решение: $\hat{\beta}_0 = -2,4899$; $\hat{\beta}_1 = 0,2539$; $\hat{\beta}_2 = -0,0014$, а уравнение регрессии примет вид

$$\bar{y}(x) = 2,5851 + 0,0004x - 0,0000002x^2$$

На рисунке представлена диаграмма рассеяния случайных величин X и Y с нанесённой линией регрессии.

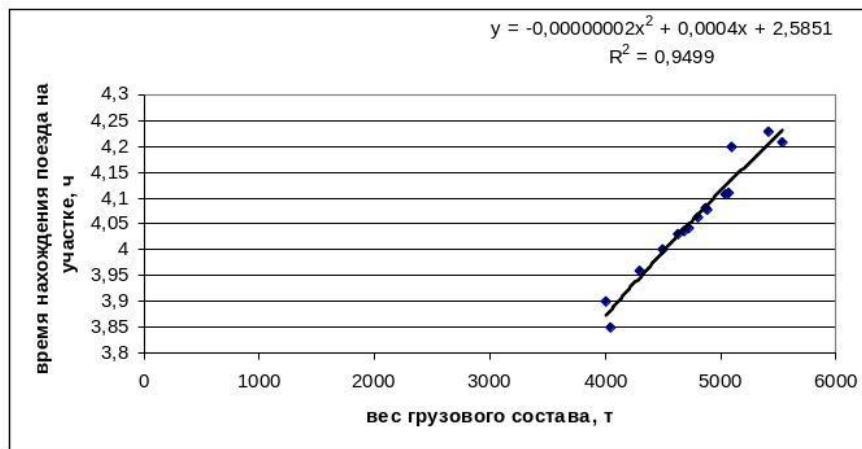


Рисунок 5 – Диаграмма рассеяния случайных величин X и Y с нанесённой линией регрессии

Оценим качество описания зависимости между величиной временем нахождения поезда на участке (Y) и весом грузового состава (X) полученным уравнением регрессии с помощью коэффициента детерминации, где – значение времени нахождения поезда на участке, предсказываемое уравнением регрессии, при среднем весе грузового

состава $\bar{x}$; $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{60,899}{15} = 4,060$ час. – среднеарифметическое наблюдаемых значений времени нахождения поезда на участке.

Таблица 5 – Значения времени нахождения поезда на участке

Вес грузового состава, т,	Время нахождения поезда на участке, час.,	Значение, предсказываемое уравнением регрессии	$(y_i - \bar{y})^2$	$(\hat{y}_i - \bar{y})^2$
5100,58	4,2	4,105	0,0196	0,00203
4885,41	4,078	4,0619	0,0003	3,9E-06
5416,94	4,23	4,165	0,0289	0,01104
4496,66	4,001	3,9794	0,0035	0,00649

4722,08	4,044	4,028	0,0003	0,00102
5537,91	4,208	4,1869	0,0219	0,01612
5074,01	4,11	4,0998	0,0025	0,00159
4807,09	4,062	4,0458	4E-06	0,0002
4046,02	3,85	3,8761	0,0441	0,03379
4683,93	4,037	4,0199	0,0005	0,0016
4872,42	4,08	4,0593	0,0004	4,6E-07
4003,22	3,9	3,8659	0,0256	0,03766
4628,01	4,03	4,0079	0,0009	0,0027
4293,44	3,96	3,9338	0,01	0,01591
5035,7	4,109	4,0922	0,0024	0,00104
Итого 71603,42	60,899	60,527	0,1609	0,152839

$$\hat{R}^2 = \frac{\sum_{i=1}^n (\hat{y}(x_i) - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 0,152839 / 0,1609 = 0,9499$$

Расчётное значение коэффициента детерминации $\hat{R}^2 = 0,9499$ указывает на удовлетворительность описания зависимости между величиной веса грузового состава (X) и времени нахождения поезда на участке (Y), выбранным уравнением регрессии. Проверим, однако, значимость оценки коэффициента детерминации с помощью статистики Фишера

Проверка значимости $\hat{R}^2$. При выполнении процедуры проверки значимости оценки коэффициента детерминации выдвигается нулевая гипотеза о том, что предложенное уравнение регрессии никак не отражает реальную зависимость между с. в., т. е. $H_0: R^2 = 0$. Альтернативная гипотеза заключается в том, что выбранная модель зависимости (уравнение регрессии) $\bar{y}(x_i)$ в достаточной степени объясняет действительную зависимость между случайными величинами, т. е. $H_a: R^2 > 0$.

Для проверки значимости оценки коэффициента детерминации используется статистика

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

имеющая F -распределение Фишера с $\square_1 = m - 1$ и $\square_2 = n - m$ степенями свободы. Здесь m – число неизвестных параметров предполагаемого уравнения регрессии. Значение статистики $\hat{F}$ сравнивается с критическим значением $F_{\square_1, \square_2, \square}$, найденным по таблицам квантилей распределения Фишера (приложение Е) при заданном уровне значимости и соответствующем числе степеней свободы. Если $F > F_{\square_1, \square_2, \square}$, то нулевая гипотеза отклоняется, вычисленный коэффициент детерминации значимо отличается от нуля, и с вероятностью ошибки $\square$ можно утверждать, что между исследуемыми величинами существует зависимость предложенного вида, и полученное уравнение регрессии может использоваться в дальнейших исследованиях.

$$\hat{F} = 0,9499 \cdot \frac{15 - 3}{(3 - 1)(1 - 0,9499)} = 113,761.$$

Вывод. Критическое значение статистики Фишера для степеней свободы $\square_1 = 3 - 1 = 2$ и $\square_2 = 15 - 3 = 12$ и уровня значимости $\square = 0,05$ составляет $F_{0,05;2;12} = 3,885$. Поскольку расчётное значение статистики Фишера больше критического ($\hat{F} = 113,761 > 3,885 = F_{0,05;2;12}$), то вычисленный коэффициент детерминации значимо отличается от нуля, и выбранное уравнение регрессионной зависимости между величинами скорости и временем нахождения поезда на участке.

Например, при весе грузового состава 6500 т можно ожидать в среднем время нахождения поезда на участке $\bar{y}(x) = 2,5851 + 0,0004 \cdot 6500 - 0,0000002 \cdot 6500^2 = 4,3401$ час.

4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций.

Основными формами текущего контроля являются устные опросы, тестирование, подготовка докладов.

Распределение баллов по видам работ

Вид работы	Кол-во баллов (максимальное значение)
Доклад	до 20 баллов
Тест	до 20 баллов
Устный опрос	до 10 баллов
Зачет	до 20 баллов

Максимальное количество баллов, которое может набрать студент в течение семестра за текущий контроль, равняется 80 баллам.

Максимальная сумма баллов, которые студент может получить на зачете, равняется 20 баллам.

Требования к проведению зачета

Зачет проводится по результатам выполнения всех видов учебной работы, предусмотренных рабочей программой дисциплины, при этом учитываются результаты текущего контроля успеваемости в течение семестра.

Критерии оценивания знаний на зачёте

Критерии оценивания	Баллы
Полно раскрыто содержание материала в объеме программы; четко и правильно даны определения и раскрыто содержание понятий; установлены причинно-следственные связи; верно использованы научные термины; для доказательства использованы различные умения, выводы из наблюдений и опытов; ответ самостоятельный, использованы ранее приобретенные знания.	15-20
Раскрыто основное содержание материала; в основном правильно даны определения понятий и использованы научные термины; определения понятий неполные, допущены незначительные нарушения последовательности изложения, небольшие неточности при использовании	10-14

научных терминов или в выводах и обобщениях из наблюдений и опытов, исправленные с помощью преподавателя.	
Усвоено основное содержание учебного материала, но изложено фрагментарно, не всегда последовательно; определения понятий недостаточно четкие; не использованы в качестве доказательства выводы и обобщения из наблюдений и опытов или допущены ошибки при их изложении; допущены ошибки и неточности в использовании научной терминологии, определении понятий, исправленные с помощью преподавателя.	5-9
Основное содержание вопроса не раскрыто; допущены грубые ошибки в определении понятий, при использовании терминологии; дополнительные и уточняющие вопросы преподавателя не приводят к коррекции ответа.	0-4

Итоговая шкала оценивания результатов освоения дисциплины

Итоговая оценка по дисциплине выставляется по приведенной ниже шкале. При выставлении итоговой оценки преподавателем учитывается работа обучающегося в течение освоения дисциплины, а также оценка по промежуточной аттестации

Баллы, полученные по текущему контролю и промежуточной аттестации	Оценка в традиционной системе
81-100	зачтено
61-80	зачтено
41-60	зачтено
0-40	Не зачтено