Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых»
(ВлГУ)

УТВЕРЖДАЮ
Проректор
по образовательной деятельности
_____ А.А Панфилов
«28» 08 2019 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Методы анализа данных

Направление подготовки: **09.03.04 Программная инженерия**

Профиль/программа подготовки: **Разработка программно-информационных систем**

Уровень высшего образования: **бакалавриат**

Форма обучения: **очная**

Семестр	Трудоемкость зач. ед./час.	Лек-ций, час.	Практич. занятий, час.	Лаборат. работ, час.	СРС, час.	Форма промежуточной аттестации (экз./зачет)
3	6/ 216	36	18	18	117	экзамен, 27
Итого	6/ 216	36	18	18	117	экзамен, 27

Владимир, 2019

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель освоения дисциплины: ознакомление студентов с методами анализа данных.
Задачи: освоение статистических методов анализа данных и принятия решений.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП ВО

Дисциплина «Методы анализа данных» относится к части учебного плана, формируемой участниками образовательных отношений.

Пререквизиты дисциплины: Математика, Алгоритмы и структуры данных, Теоретические основы дискретных вычислений.

3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ДИСЦИПЛИНЫ

Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения ОПОП

Код формируемых компетенций	Уровень освоения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций (показатели освоения компетенции)
1	2	3
ОПК-1	Частичное освоение	Знать: основы математики, вычислительной техники, программирования. Уметь: решать стандартные профессиональные задачи с применением естественнонаучных и общеинженерных знаний, методов математического анализа и моделирования. Иметь навыки: теоретического и экспериментального исследования объектов профессиональной деятельности.
ПК-3	Частичное освоение	Знать: Возможности современных и перспективных средств разработки программных продуктов, технических средств Методологии разработки программного обеспечения и технологии программирования Методологии и технологии проектирования и использования баз данных Языки формализации функциональных спецификаций Методы и средства проектирования программного обеспечения Методы и средства проектирования программных интерфейсов Методы и средства проектирования баз данных Принципы построения архитектуры программного обеспечения и виды архитектуры программного обеспечения Типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке программного обеспечения Уметь: Вырабатывать варианты реализации программного обеспечения Проводить оценку и обоснование рекомендуемых решений Использовать существующие типовые решения и шаблоны проектирования программного обеспечения Применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов Иметь навыки: Анализа требований к программному обеспечению Разработки технических спецификаций на программные компоненты и их взаимодействие Проектирования программного обеспечения

4. ОБЪЕМ И СТРУКТУРА ДИСЦИПЛИНЫ

Трудоемкость дисциплины составляет 6 зачетных единиц, 216 часов.

№ п/п	Наименование разделов/тем дисциплины	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)				Объем учебной работы с применением интерактивных методов (в часах / %)	Формы текущего контроля успеваемости форма промежуточной аттестации
				Лекции	Лабораторные работы	Практические занятия	СРС		
1	2	3	4	5	6	7	8	9	10
1	Введение. Методы анализа данных	3	1	2	2		6	2 /50	Рейтинг-контроль №1
2	Случайные события в системах. Вероятностное описание событий	3	2	2	2		6	2 /50	
3	Случайные величины. Распределение дискретных величин	3	3	2	2		6	2 /50	
4	Случайные величины. Распределение непрерывных величин	3	4	2	2		6	2 /50	
5	Многомерные распределения. Закон больших чисел. Центральная предельная теорема.	3	5	2	2		6	2 /50	
6	Статистическая оценка параметров распределения.	3	6	2	2		6	2 /50	
7	Статистическая проверка гипотез	3	7	2	2		6	2 /50	
8	Случайные процессы и их характеристики. Аналитическое описание случайных процессов.	3	8	2	2		6	2 /50	
9	Основы теории корреляционного анализа. Парная корреляция, оценка тесноты связи	3	9	2	2		6	2 /50	
10	Множественная корреляция, оценка тесноты связи	3	10	2		2	7	2 /50	Рейтинг-контроль №2
11	Основы дисперсионного анализа.	3	11	2		2	7	2 /50	

№ п/п	Наименование разделов/тем дисциплины	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)				Объем учебной работы с применением интерактивных методов (в часах / %)	Формы текущего контроля успеваемости форма промежуточной аттестации
				Лекции	Лабораторные работы	Практические занятия	СРС		
1	2	3	4	5	6	7	8	9	10
	Однофакторный и многофакторный дисперсионные анализы								Рейтинг-контроль №3
12	Регрессионный анализ. Множественная регрессия. Оценка качества модели регрессии.	3	13	2		2	7	2 /50	
13	Нелинейная регрессия	3	14	2		2	7	2 /50	
14	Системы линейных регрессионных уравнений	3		2		2	7	2 /50	
15	Компонентный анализ. . Метод главных компонент.	3	15	2		2	7	2 /50	
16	Методы анализа больших систем, факторный анализ	3	16	2		2	7	2 /50	
17	Временные ряды. Анализ временных рядов	3	17	2		2	7	2 /50	
18	Заключение. Прогнозирование временных рядов	3	18	2		2	7	2 /50	
Всего за семестр:				36	18	18	117	36 /50	Экзамен, 27
Наличие в дисциплине КП/КР						нет			
Итого по дисциплине				36	18	18	117	36/50	Экзамен, 27

Содержание лекционных занятий по дисциплине

1. Введение. Методы анализа данных.
1.1. Понятие интеллектуального анализа данных
1.2. Набор данных и их атрибутов
1.3. Задачи *Data Mining*
1.4. Основы анализа данных
1.5. Основные этапы интеллектуального анализа
1.6. Инструментальные средства анализа данных
2. Случайные события в системах. Вероятностное описание событий
2.1. Испытание. Поле событий. Операции над событиями

2.2. Частость и вероятность
2.3. Основные аксиомы теории вероятностей
2.4. Случайные события
3. Случайные величины. Распределение дискретных величин
3.1. Определение случайной величины
3.2. Законы распределения дискретных случайных величин
3.3. Биноминальное распределение
3.4. Закон распределения Пуассона
4. Случайные величины. Распределение непрерывных величин
4.1. Экспоненциальное распределение
4.2. Равномерное распределение
4.3. Нормальное распределение
4.4. Распределение χ - квадрат
4.5. Распределение Стьюдента
4.6. Распределение Фишера
5. Многомерные распределения. Закон больших чисел. Центральная предельная теорема
5.1. Двумерное дискретное распределение
5.2. Двумерное непрерывное нормальное распределение
5.3. Многомерное распределение
5.4. Закон больших чисел
5.5. Основные предельные законы теории вероятностей
6. Статистическая оценка параметров распределения.
6.1. Статистическое распределение выборки
6.2. Точечные оценки параметров генеральной совокупности по выборке
6.3. Интервальное оценивание
7. Статистическая проверка гипотез
7.1. Уровень значимости и мощность критерия. Ошибки при проверке гипотез
7.2. Общая схема проверки гипотез
7.3. Односторонние и двусторонние критерии
7.4. Некоторые типичные задачи проверки параметрических гипотез
7.5. Непараметрические гипотезы. Критерии согласия Пирсона и Колмогорова
7.6. Критерий Пирсона
8. Случайные процессы и их характеристики. Аналитическое описание случайных процессов.
8.1. Функциональные характеристики случайного процесса
8.2. Одномерная функция распределения вероятностей
8.3. Одномерная плотность вероятностей
8.4. Математическое ожидание. Функция дисперсии
8.5. Корреляционные функции случайных процессов
8.6. Свойства функций автоковариации и автокорреляции
8.7. Взаимные моменты случайных процессов второго порядка
8.8. Классификация случайных процессов
9. Основы теории корреляционного анализа. Парная корреляция, оценка тесноты связи
9.1. Функциональные и корреляционные связи между переменными
9.2. Корреляционный анализ
9.3. Парный коэффициент корреляции, основные свойства
9.4. Проверка значимости парного коэффициента корреляции
10. Множественная корреляция, оценка тесноты связи
10.1. Задачи многомерного корреляционного анализа
10.2. Множественный коэффициент корреляции

10.3. Частный коэффициент корреляции
11. Основы дисперсионного анализа. Однофакторный и многофакторный дисперсионные анализы
11.1. Формулировка гипотез в дисперсионном анализе
11.2. Однофакторный дисперсионный анализ
11.3. Двухфакторный дисперсионный анализ
12. Регрессионный анализ. Множественная регрессия. Оценка качества модели регрессии
12.1. Формулирование задача регрессионного анализа
12.2. Вычисление коэффициентов регрессии
12.3. Статистический анализ уравнения регрессии
12.4. Проверка выполнения предпосылок МНК
12.5. Оценка влияния отдельных факторов на зависимую переменную на основе модели
12.6. Использование многофакторных моделей для анализа и прогнозирования развития явлений и процессов
13. Нелинейная регрессия
13.1. Виды моделей нелинейной регрессии
13.2. Нелинейная модель относительно включенных в структуру моделей влияющих факторов
13.3. Нелинейная модель по оцениваемым параметрам
13.4. Примеры нелинейных моделей, встречаемые экономике
14. Системы линейных регрессионных уравнений
14.1. Виды систем линейных регрессионных уравнений
14.2. Условия идентифицируемости модели
14.3. Косвенный метод наименьших квадратов при оценивании коэффициентов структурной модели
15. Компонентный анализ. Метод главных компонент
15.1. Идея метода компонентного анализа
15.2. Метод главных компонент
15.3. Принципы выбора числа главных компонент
15.4. Построение регрессионного уравнения на главных компонентах
16. Методы анализа больших систем, факторный анализ
16.1. Подходы к решению задач, в которых используются статистические данные
16.2. Факторный анализ. Сущность факторного анализа
16.3. Условия применения факторного анализа
16.4. Последовательность факторного анализа
17. Временные ряды. Анализ временных рядов
17.1. Модели временных рядов. Аддитивная модель
17.2. Структура временных рядов, выделение компонент структуры
17.3. Статистические оценки взаимосвязи двух временных рядов
18. Заключение. Прогнозирование временных рядов
18.1. Основное содержание прогнозирования процессов
18.2. Методы прогнозирования временных рядов
18.3. Оценка адекватности и точности трендовых моделей прогноза

Содержание лабораторных занятий по дисциплине

Лабораторная работа №1. Описательная статистика. Опции диалогового окна Описательная статистика (2ч)
Лабораторная работа №2. Случайные события. Вероятностное описание событий (2ч)
Лабораторная работа №3. Генерация случайных чисел. Распределение дискретных величин (2ч)
Лабораторная работа №4. Генерация непрерывных случайных величин (2ч)
Лабораторная работа №5. Гистограмма. Опции диалогового окна Гистограмма (2ч)

Лабораторная работа №6. Выборка. Опции диалогового окна Выборка (2ч)
Лабораторная работа №7. Статистическая проверка гипотез (2ч)
Лабораторная работа №8. Корреляционный анализ многомерной выборки (2ч)
Лабораторная работа №9. Анализ Фурье (2ч)

Содержание практических занятий по дисциплине

1. Однофакторный дисперсионный анализ.
2. Двухфакторный дисперсионный анализ.
3. Регрессионный анализ. Множественная регрессия.
4. Нелинейная регрессия.
5. Системы линейных регрессионных уравнений.
6. Компонентный анализ. Метод главных компонент.
7. Методы анализа больших систем, факторный анализ.
8. Временные ряды. Анализ временных рядов.
9. Прогнозирование временных рядов.

5.ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

В преподавании дисциплины «Методы анализа данных» используются разнообразные образовательные технологии как традиционные, так и с применением активных и интерактивных методов обучения.

Активные и интерактивные методы обучения:

- интерактивные лекции с мультимедийным комплектом слайдов (темы № 1 – 18);
- разбор конкретных ситуаций (темы № 9 – 18);
- выполнение индивидуального лабораторного задания (темы № 1 – 9).

6. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ И УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

Перечень контрольных вопросов и заданий для проведения текущего контроля:

3 семестр:

Рейтинг-контроль № 1

1. Что называется испытанием и событием в испытании?
2. Какие бывают события в испытаниях?
3. Что называется частостью и вероятностью событий?
4. Как влияет объем выборки на частость?
5. Назовите виды случайных величин, их особенности.
6. Способы задания случайных величин.
7. Числовые характеристики случайной величины.
8. Как вычисляются числовые характеристики выборочного распределения?
9. Что называется доверительным интервалом параметра?
10. Какие требования предъявляются к оценкам параметров модели

Рейтинг-контроль № 2

1. Чем характеризуется случайный процесс.
2. Одномерный закон распределения мгновенных значений случайной функции и связанные с ним основные характеристики.
3. Гауссовский случайный процесс. Центральная предельная теорема.
4. Стационарные случайные процессы и их характеристики.
5. Основные задачи корреляционного анализа.
6. Оценка выборочного коэффициента парной корреляции по статистическим данным

7. Дисперсионный анализ и его назначение. Основная идея дисперсионного анализа.

Рейтинг-контроль № 3

1. Какой априорной информацией располагает экспериментатор приступая к поиску математической модели.

2. Оценка влияния отдельных факторов на зависимую переменную на основе модели.

3. Использование многофакторных моделей для анализа и прогнозирования развития систем и процессов.

4. В чем идея метода компонентного анализа?

5. Когда имеет смысл проводить факторный анализ?

6. Запишите модель аддитивного случайного процесса. Приведите интерпретацию компонентов модели.

7. Что понимается под коинтеграцией причинно следственной зависимости в уровнях двух или более временных рядов?

8. Какие предположения лежат в основе алгоритмов прогнозирования?

9. От чего зависит надежность и точность прогноза?

Перечень вопросов к экзамену (аттестация по итогам освоения дисциплины)

1. Случайные события в системах. Операции над событиями. Частость и вероятность. Основные аксиомы теории вероятностей.

2. Определения случайной величины. Законы распределения дискретных случайных величин.

3. Законы распределения непрерывных случайных величин. Нормальный закон распределения.

4. Закон больших чисел. Основные предельные законы теории вероятностей.

5. Проверка статистических гипотез. Критерий значимости при нормальном распределении.

6. Случайные величины и их характеристики. Оценка характеристик.

7. Случайные процессы и их характеристики. Оценка характеристик.

8. Парная корреляция. Оценка тесноты корреляционных связей.

9. Множественная корреляция. Оценка тесноты корреляционных связей и их значимости.

10. Однофакторный дисперсионный анализ. Интерпретация результатов анализа.

11. Двухфакторный дисперсионный анализы. Интерпретация результатов анализа.

12. Парная регрессия. Оценка параметров и качества регрессионных моделей.

13. Множественная регрессия. Оценка параметров модели.

14. Множественная регрессия. Оценка качества регрессионных моделей. Точечный и интервальный прогнозы.

15. Нелинейная модель относительно включенных в структуру моделей влияющих факторов

16. Нелинейная модель по оцениваемым параметрам

17. Виды систем линейных регрессионных уравнений

18. Косвенный метод наименьших квадратов при оценивании коэффициентов структурной модели

19. Метод главных компонент. Линейная модель метода главных компонент. Решаемые задачи.

20. Факторный анализ, решаемые задачи.

21. Модели временных рядов. Аддитивный случайный процесс. Оценка параметров модели и ее адекватности.

22. Статистические оценки взаимосвязи двух временных рядов. Коинтеграция временных рядов.

23. Методы и алгоритмы прогнозирования временных рядов. Оценка точности прогноза временного ряда.

Перечень заданий для самостоятельной работы студентов

3 семестр:

Примерный перечень практических заданий

1. Однофакторный дисперсионный анализ.
2. Двухфакторный дисперсионный анализ.
3. Регрессионный анализ. Множественная регрессия.
4. Нелинейная регрессия по оцениваемым параметрам.
5. Системы линейных регрессионных уравнений.
6. Компонентный анализ. Метод главных компонент.
7. Методы анализа больших систем, факторный анализ.
8. Временные ряды. Анализ временных рядов.
9. Прогнозирование временных рядов.

Самостоятельная работа обучающихся заключается в самостоятельном изучении отдельных тем, практической реализации типовых заданий по этим темам. Контроль выполнения самостоятельной работы проводится при текущих контрольных мероприятиях. Учебно-методическое обеспечение самостоятельной работы – основная литература [1–4], дополнительная литература [1-2].

Фонд оценочных средств для проведения аттестации уровня сформированности компетенций обучающихся по дисциплине оформляется отдельным документом

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

7.1. Книгообеспеченность

Наименование литературы: автор, название, вид издания, издательство	Год издания	КНИГООБЕСПЕЧЕННОСТЬ	
		Количество экземпляров изданий в библиотеке ВлГУ в соответствие с ФГОС ВО	Наличие в электронной библиотеке ВлГУ
1	2	3	4
Основная литература			
1. Гулай, Т.А. Теория вероятностей и математическая статистика [Электронный ресурс] : учебное пособие / Т.А. Гулай, А.Ф. Долгополова, Д.Б. Литвин, С.В. Мелешко. - 2-е изд., доп. – Ставрополь: АГРУС, 2013. - 260 с.	2013		http://znanium.com/catalog.php?bookinfo=514780
2. Эконометрика : учебник для вузов по специальности 061700 "Статистика" / И. И. Елисеева [и др.] ; под ред. И. И. Елисеевой .— Москва : Финансы и статистика, 2002 .— 342 с.	2002	6	
3. Дубров А.М., Мхитарян В.С.,	2003		https://institutiones.com/d

Трошин Л.И. Многомерные статистические методы; Учебник. - М.: Финансы и статистика, 2003.-352 с.			ownload/books/1781-mnogomernye-statisticheskie-metody-dubrov.html
4. Лекции по теории вероятностей и математической статистике [Электронный ресурс]: учебник / Прохоров Ю.В., Пономаренко Л.С. М.: Издательство Московского государственного университета, 2012. (Классический университетский учебник).	2012		http://www.iprbookshop.ru/13173.html.— ЭБС «IPRbooks».
Дополнительная литература			
1. Макаров, Руслан Ильич. Основы планирования и обработки экспериментальных данных [Электронный ресурс]: курс лекций для студентов ВлГУ, обучающихся по направлению 231000.62 "Программная инженерия". Профиль подготовки «Разработка программно-информационных систем»: учебное электронное издание / Владимир: Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых (ВлГУ), 2014 .— 180 с	2014	1	http://e.lib.vlsu.ru/bitstream/123456789/3646/1/00474.doc
2. Макаров, Руслан Ильич. Основы планирования и обработки экспериментальных данных [Электронный ресурс]: методические указания к лабораторным занятиям для студентов ВлГУ, обучающихся по направлению 231000.62 "Программная инженерия". Профиль подготовки «Разработка программно-информационных систем» : учебное электронное издание / Владимир : Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых (ВлГУ), 2014 .— 37 с.	2014	1	http://e.lib.vlsu.ru/bitstream/123456789/3647/1/00475.doc

7.2. Периодические издания

1. Вестник компьютерных и информационных технологий *ISSN* 1810-7206.
2. Современные наукоёмкие технологии *ISSN* 1812-7320.

7.3. Интернет-ресурсы

1. *www.edu.ru* – портал российского образования
2. *www.elbib.ru* – портал российских электронных библиотек
3. *www.eLibrary.ru* – научная электронная библиотека
4. *www.intuit.ru* - интернет университета информационных технологий
5. *library.vlsu.ru* - научная библиотека ВлГУ
6. *https://vlsu.bibliotech.ru/* - электронная библиотечная система ВлГУ

8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Для реализации данной дисциплины имеются специальные помещения для проведения занятий лекционного типа, занятий лабораторного типа, текущего контроля и промежуточной аттестации.

Лабораторные работы проводятся в аудитории 404а - 2.

Перечень используемого лицензионного программного обеспечения:

1. *Windows* 10 Корпоративная *MSDN* подписка: Идентификатор подписчика:700619248
2. *Microsoft Office* 2013 *Microsoft Open License* 66772217
3. *MATLAB R*2010*b License Number*: 357594
4. *MathCAD* 14.0 *M*011 (14.0.1.286 [709051735]) Лицензия: *PKG*-7518-*FN*

Рабочую программу составила: д.т.н., проф. Макаров Р.И.

Рецензент: к.т.н., генеральный директор ООО «Системный подход» Шориков А.В.

Программа рассмотрена и одобрена на заседании кафедры ИСПИ
протокол № 1 от 28.08.19 года.
Заведующий кафедрой _____ Жигалов И.Е.

Рабочая программа рассмотрена и одобрена на заседании учебно-методической комиссии направления 09.03.04 - Программная инженерия
протокол № 1 от 28.08.19 года.
Председатель комиссии _____ И.Е. Жигалов

ЛИСТ ПЕРЕУТВЕРЖДЕНИЯ
РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ

Рабочая программа одобрена на 2020/21 учебный год
Протокол заседания кафедры № 1 от 31.08.20 года
Заведующий кафедрой _____

Рабочая программа одобрена на _____ учебный год
Протокол заседания кафедры № _____ от _____ года
Заведующий кафедрой _____

Рабочая программа одобрена на _____ учебный год
Протокол заседания кафедры № _____ от _____ года
Заведующий кафедрой _____

Рабочая программа одобрена на _____ учебный год
Протокол заседания кафедры № _____ от _____ года
Заведующий кафедрой _____

Рабочая программа одобрена на _____ учебный год
Протокол заседания кафедры № _____ от _____ года
Заведующий кафедрой _____

Рабочая программа одобрена на _____ учебный год
Протокол заседания кафедры № _____ от _____ года
Заведующий кафедрой _____

ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ

в рабочую программу дисциплины «Методы анализа данных»
образовательной программы направления подготовки 09.03.04 «Программная инженерия»,
направленность «Разработка программно-информационных систем» (уровень бакалавриат)

Номер изменения	Внесены изменения в части/разделы рабочей программы	Исполнитель ФИО	Основание (номер и дата протокола заседания кафедры)
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Зав. кафедрой _____ / _____
Подпись *ФИО*