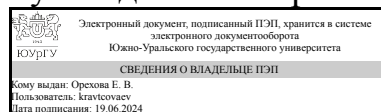


ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ:
Руководитель направления



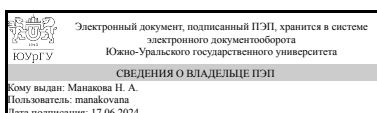
Е. В. Орехова

РАБОЧАЯ ПРОГРАММА

дисциплины 1.Ф.14.М2.01 Анализ данных и технологии работы с данными
для направления 45.03.03 Фундаментальная и прикладная лингвистика
уровень Бакалавриат
форма обучения очная
кафедра-разработчик Уравнения математической физики

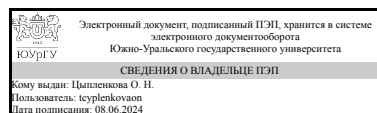
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 45.03.03 Фундаментальная и прикладная лингвистика, утверждённым приказом Минобрнауки от 24.04.2018 № 323

Зав.кафедрой разработчика,
д.физ.-мат.н., проф.



Н. А. Манакова

Разработчик программы,
к.физ.-мат.н., доц., доцент



О. Н. Цыпленкова

1. Цели и задачи дисциплины

Целью освоения дисциплины «Анализ данных и технологии работы с данными» является формирование теоретических и практических компетенций в области целостного представления, понимания места и роли, а также применения методов статистического анализа данных, а также обучение студентов современным программным средствам, в которых реализованы модули, осуществляющие решение задач анализа данных. Задачи дисциплины: - дать теоретические основы по спектру наиболее распространенных статистических методов анализа данных и условий их применения; - дать основы количественных методов оценки адекватности и точности построенных моделей; - привить навыки и умения практического применения компьютерных технологий при анализе и прогнозировании социально-экономических показателей (построение линейных и нелинейных моделей прогнозирования на основе регрессионного анализа, оценка их параметров, расчёт всех необходимых статистик для анализа моделей); -изучение концепции и технологии современного анализа данных на компьютере; -изучение принципов работы программных средств, предназначенных для статистического анализа данных; -изучение современных визуальных методов анализа данных и использования их для статистического вывода и формулировки гипотез о структуре данных; -выработка умения самостоятельного решения задач по выбору методов анализа в практических ситуациях; -получение навыков применения программных систем, предназначенных для статистического анализа данных, а также тестировании программных модулей на модельных данных; - обеспечить бакалаврам прочное и осознанное желание владеть системой компьютерный анализ данных, способствующей их профессиональной успешности и востребованности на рынке труда.

Краткое содержание дисциплины

Дисциплина «Анализ данных и технологии работы с данными» знакомит студентов с основами статистических методов анализа данных, а также с программными продуктами, осуществляющие работу с данными.

2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
УК-2 Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений	Знает: способы сбора, обработки и анализа данных для решения своих профессиональных задач с учётом имеющихся ресурсов и правовых норм Умеет: применять математические методы обработки данных для выбора и реализации оптимального способа решения

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
---	---

Нет	1.Ф.14.М2.02 Программирование для анализа данных, 1.Ф.14.М2.03 Приложения и практика анализа данных
-----	--

Требования к «входным» знаниям, умениям, навыкам студента, необходимым при освоении данной дисциплины и приобретенным в результате освоения предшествующих дисциплин:

Нет

4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 4 з.е., 144 ч., 72,5 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах
		Номер семестра
		3
Общая трудоёмкость дисциплины	144	144
<i>Аудиторные занятия:</i>	64	64
Лекции (Л)	32	32
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	32	32
Лабораторные работы (ЛР)	0	0
<i>Самостоятельная работа (СРС)</i>	71,5	71,5
Подготовка к контрольным точкам	25	25
Подготовка к дифференцированному зачету	21,5	21.5
Выполнение индивидуальных домашних заданий	25	25
Консультации и промежуточная аттестация	8,5	8,5
Вид контроля (зачет, диф.зачет, экзамен)	-	диф.зачет

5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Случайные величины	12	6	6	0
2	Первичная обработка выборки (данных)	10	4	6	0
3	Описательная статистика (вычисление числовых характеристик)	4	2	2	0
4	Проверка статистических гипотез	12	6	6	0
5	Корреляционный анализ	10	6	4	0
6	Регрессионный анализ	16	8	8	0

5.1. Лекции

№	№	Наименование или краткое содержание лекционного занятия	Кол-
---	---	---	------

лекции	раздела		во часов
1-2	1	Дискретные случайные величины	3
2-3	1	Непрерывные случайные величины	3
4	2	Построение статистического распределения. Построение гистограммы относительных частот.	4
5	3	Вычисление точечных оценок математического ожидания и дисперсии.	2
6	4	Проверка гипотезы о нормальном распределении случайной величины	2
7	4	Классические методы статистики: t-критерий Стьюдента	2
8	4	Двухвыборочный тест для дисперсии: F-тест в EXCEL	2
9	5	Корреляция	2
10	5	Значимость коэффициента корреляции, доверительный интервал	2
11	5	Оценка значимости статистических характеристик корреляции	2
12	6	Линейная регрессия	2
13	6	Проверка гипотез при использовании линейной регрессии	2
14	6	Множественная линейная регрессия	2
15	6	Полиномиальная регрессия. Многофакторный дисперсионный анализ	2

5.2. Практические занятия, семинары

№ занятия	№ раздела	Наименование или краткое содержание практического занятия, семинара	Кол-во часов
1-2	1	Дискретные случайные величины	3
2-3	1	Непрерывные случайные величины	3
4-5	2	Первичная обработка данных. Построение вариационного и статистического ряда, гистограммы, полигона.	3
4	2	Контрольная работа на тему: "Случайные величины"	1
6	2	Построение статистического ряда и гистограммы с помощью Пакета Анализа	2
7	3	Описательная статистика (вычисление числовых характеристик)	2
8	4	Использование таблиц критических значений и р-значений. Проверка гипотезы о нормальном распределении	2
9	4	Проверка гипотезы о равенстве средних — Т-тесты, проверка гипотезы о равенстве дисперсий — F-тесты	2
10	4	Контрольная работа по теме: "Проверка статистических гипотез"	2
11	5	Диаграмма рассеяния. Коэффициент линейной корреляции Пирсона	2
12	5	Проверка значимости коэффициента регрессии	2
13	6	Линейная регрессия. Построение уравнения парной регрессии с помощью линии тренда точечной диаграммы	2
14	6	Линейная регрессия – многомерный случай. Оценка качества уравнения регрессии, проверка значимости уравнения и коэффициентов	2
15	6	Полиномиальная регрессия. Другие виды нелинейной регрессии	2
16	6	Контрольная работа по теме "Корреляционный и регрессионный анализ"	2

5.3. Лабораторные работы

Не предусмотрены

5.4. Самостоятельная работа студента

Выполнение СРС			
Подвид СРС	Список литературы (с указанием разделов, глав, страниц) / ссылка на ресурс	Семестр	Кол-во часов
Подготовка к контрольным точкам	ПУМД, осн. лит. 1 (гл. 40-45), 2 (с.151-289), 3 (с.187-361); ПУМД, доп. лит. 1 (гл. 3,4, 8-13); ЭУМД основная (п. 1, п. 3)	3	25
Подготовка к дифференцированному зачету	ПУМД, осн. лит. 1 (гл. 40-45), 2 (с.151-289), 3 (с.187-361); ЭУМД основная (п. 1, п. 3)	3	21,5
Выполнение индивидуальных домашних заданий	ПУМД, осн. лит. 1 (гл. 40-45), 2 (с.151-289), 3 (с.187-361); Метод. пособ.	3	25

6. Фонд оценочных средств для проведения текущего контроля успеваемости, промежуточной аттестации

Контроль качества освоения образовательной программы осуществляется в соответствии с Положением о балльно-рейтинговой системе оценивания результатов учебной деятельности обучающихся.

6.1. Контрольные мероприятия (КМ)

№ КМ	Се-местр	Вид контроля	Название контрольного мероприятия	Вес	Макс. балл	Порядок начисления баллов	Учитывается в ПА
1	3	Текущий контроль	Контрольная работа №1	0,15	15	Продолжительность – 2 академических часа. Студент должен самостоятельно решить задачи, оформить их решение на отдельном листе. Контрольная работа состоит из 2 задач. Максимальный балл за решение первой задачи – 7 баллов, за решение второй – 8 баллов. Первая задача оценивается следующим образом: 1) 2 балла - верно составлен закон распределения дискретной случайной величины, 1 балл – выбран верный метод решения, есть 1–2 ошибки; 0 баллов – отсутствует решение или сделано более 2	дифференцированный зачет

						грубых ошибок. 2) 3 балла – вычислены три числовые характеристики; 2 балла – вычислены две числовые характеристики; 1 балл – вычислена одна числовая характеристика; 0 баллов – отсутствует решение или сделано более 2 грубых ошибок. 3) 1 балл – правильно начерчен многоугольник распределения; 0 баллов – отсутствует решение или сделано более 2 грубых ошибок. 4) 1 балл – правильно начерчен многоугольник распределения; и график функции распределения (1 балл). 0 баллов – отсутствует решение или сделано более 2 грубых ошибок. Вторая задача состоит из 4 вопросов, каждый из которых оценивается следующим образом: 2 балла – задача решена верно; 1 балл – выбран верный метод решения задачи, возможна арифметическая ошибка; 0 баллов – отсутствует решение или сделано более 2 грубых ошибок.	
2	3	Текущий контроль	Контрольная работа №2	0,15	15	Продолжительность – 2 академических часа. Студент должен самостоятельно	дифференцированный зачет

						решить задачи, оформить их решение на отдельном листе. Контрольная работа состоит из 5 задач. Максимальный балл за решение задачи – 3 балла. Каждая задача оценивается следующим образом: 3 балла – задача решена верно, ошибок нет; 2 балла – выбран верный метод решения задачи, возможна арифметическая ошибка; 1 балл – выбран верный метод решения, есть 1–2 грубые ошибки; 0 баллов – отсутствует решение или сделано более 2 грубых ошибок.	
3	3	Текущий контроль	Контрольная работа №3	0,15	15	Продолжительность – 2 академических часа. Студент должен самостоятельно решить задачи, оформить их решение на отдельном листе. Контрольная работа состоит из 5 задач. Максимальный балл за решение задачи – 3 балла. Каждая задача оценивается следующим образом: 3 балла – задача решена верно, ошибок нет; 2 балла – выбран верный метод решения задачи, возможна арифметическая ошибка; 1 балл – выбран верный метод решения, есть 1–2 грубые ошибки; 0 баллов – отсутствует решение	дифференцированный зачет

						или сделано более 2 грубых ошибок.	
4	3	Текущий контроль	Теоретическая контрольная работа №1	0,08	8	Контрольная точка Т1 проводится на лекционном занятии на последнем занятии семестра. Продолжительность – 40 минут. Работа состоит из 4 теоретических вопросов. Максимальная оценка за вопрос составляет 2 балла. При оценке используется следующая шкала: 2 балла – приведен полный ответ на вопрос, все использованные формулы верны, записаны все требуемые свойства; 1 балл – в ответе содержатся 2–3 ошибки или ответ неполный, но при этом изложено не менее 60% полного ответа; 0 баллов – изложено менее 60% верного ответа на вопрос.	дифференцированный зачет
5	3	Текущий контроль	Теоретическая контрольная работа №2	0,08	8	Контрольная точка Т2 проводится на лекционном занятии на последнем занятии семестра. Продолжительность – 20 минут. Работа состоит из 4 теоретических вопросов. Максимальная оценка за вопрос составляет 2 балла. При оценке используется следующая шкала: 2 балла – приведен полный ответ на вопрос, все использованные формулы верны,	дифференцированный зачет

						записаны все требуемые свойства; 1 балл – в ответе содержатся 2–3 ошибки или ответ неполный, но при этом изложено не менее 60% полного ответа; 0 баллов – изложено менее 60% верного ответа на вопрос.	
6	3	Текущий контроль	Проверка конспекта лекций и посещаемости	0,07	7	Максимальный балл - 7. При оценке используется следующая шкала: 7*m/n баллов – приведен полный конспект лекций, где m-количество посещенных занятий, а n - общее количество занятий.	дифференцированный зачет
7	3	Текущий контроль	Активная работа	0,04	4	Контрольная точка П1 учитывает активность студента с 1 по 5 практическое занятие и выполнение домашних заданий. На каждом практическом занятии студент может получить "плюсик" за выход к доске. Предъявляя выполненную домашнюю работу, студент также зарабатывает "плюсик". При оценке используется следующая шкала: 4 балла - у студента 8 и более "плюсиков"; 3 балла - у студента 6 или 7 "плюсиков"; 2 балла - у студента 4 или 5 "плюсиков"; 1 балл - у студента 2 или 3 "плюсика"; 0 баллов - у студента 0 или 1 "плюсика".	дифференцированный зачет
8	3	Текущий контроль	Активная работа 2	0,04	4	Контрольная точка П2 учитывает активность	дифференцированный зачет

						студента с 6 по 11 практическое занятие и выполнение домашних заданий. На каждом практическом занятии студент может получить "плюсик" за выход к доске. Предъявляя выполненную домашнюю работу, студент также зарабатывает "плюсик". При оценке используется следующая шкала: 4 балла - у студента 8 и более "плюсиков"; 3 балла - у студента 6 или 7 "плюсиков"; 2 балла - у студента 4 или 5 "плюсиков"; 1 балл - у студента 2 или 3 "плюсика"; 0 баллов - у студента 0 или 1 "плюсика".	
9	3	Текущий контроль	Активная работа 3	0,04	4	Контрольная точка ПЗ учитывает активность студента с 12 по 16 практическое занятие и выполнение домашних заданий. На каждом практическом занятии студент может получить "плюсик" за выход к доске. Предъявляя выполненную домашнюю работу, студент также зарабатывает "плюсик". При оценке используется следующая шкала: 4 балла - у студента 8 и более "плюсиков"; 3 балла - у студента 6 или 7 "плюсиков"; 2 балла - у студента 4 или 5 "плюсиков"; 1 балл - у студента 2 или 3 "плюсика";	дифференцированный зачет

						0 баллов - у студента 0 или 1 "плюсика".	
10	3	Текущий контроль	Решение индивидуальных заданий	0,2	20	Контрольная точка С служит для контроля самостоятельной работы студентов. Задание выдается студенту в начале семестра. Вариант определяется порядковым номером студента в журнале группы. Работа выполняется студентом самостоятельно вне аудитории и сдается студентом на последней неделе текущего семестра. Контрольная точка содержит 20 задач. Студент должен самостоятельно решить задачи, привести условие задачи, аккуратно оформить их подробное решение, привести в решении использованные свойства и формулы. Каждая правильно решенная задача оценивается в 1 балл.	дифференцированный зачет
11	3	Промежуточная аттестация	Дифференцированный зачет	-	40	Дифференцированный зачет проходит в письменной форме. Билет состоит из 8 заданий. Максимальный балл за каждое задание – 5 баллов. 5 баллов – задание решено верно, 4 балла – задание решено в целом правильно, содержится не более двух негрубых ошибок, не повлиявших на общий ход решения задачи, верно выбран метод решения задачи, запись решения последовательная и	дифференцированный зачет

					математически грамотная, решение доведено до ответа; 3 балла – в решении содержатся 2–3 ошибки, не повлиявшие существенно на ход решения, или решение не доведено до ответа, но при этом изложено не менее 80% полного решения, 2 балла - в решении содержатся ошибки, не повлиявшие существенно на ход решения, или решение не доведено до ответа, но при этом изложено не менее 60% полного решения, 1 балл – в процессе решения допущены существенные ошибки, показавшие, что студент не владеет обязательными знаниями и умениями по данной теме, или изложено менее 40% полного решения; 0 баллов – неверно выбран метод решения или изложено менее 20% полного решения.	
--	--	--	--	--	--	--

6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
дифференцированный зачет	При оценивании результатов учебной деятельности обучающегося по дисциплине используется балльно-рейтинговая система оценивания результатов учебной деятельности обучающихся (утверждена приказом ректора от 24.05.2019 г. № 179). Рейтинг обучающегося по дисциплине определяется только по результатам текущего контроля Ртек. Для расчета рейтинга обучающегося по дисциплине используется следующая формула: $\text{Ррейт} = \text{Ртек} + \text{б.}$ Студент вправе пройти контрольное мероприятие в рамках промежуточной аттестации для улучшения своего рейтинга и может получить оценку по дисциплине согласно п. 2.4 Положения. Контрольное мероприятие промежуточной аттестации проводится во время дифференцированного зачета в виде письменной работы.	В соответствии с пп. 2.5, 2.6 Положения

6.3. Паспорт фонда оценочных средств

Компетенции	Результаты обучения	№ КМ										
		1	2	3	4	5	6	7	8	9	10	11
УК-2	Знает: способы сбора, обработки и анализа данных для решения своих профессиональных задач с учётом имеющихся ресурсов и правовых норм	+	+	+	+	+	+	+	+	+	+	+
УК-2	Умеет: применять математические методы обработки данных для выбора и реализации оптимального способа решения	+	+	+				+	+	+	+	+

Типовые контрольные задания по каждому мероприятию находятся в приложениях.

7. Учебно-методическое и информационное обеспечение дисциплины

Печатная учебно-методическая документация

а) основная литература:

1. Вся высшая математика Т. 5 В 6 т.: Учеб. М. Л. Краснов, А. И. Киселев, Г. И. Макаренко и др. - М.: Эдиториал УРСС, 2001. - 293,[1] с.
2. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике [Текст] учеб. пособие для вузов В. Е. Гмурман. - 11-е изд., перераб. и доп. - М.: Юрайт, 2013. - 403, [1] с. ил.
3. Гмурман, В. Е. Теория вероятностей и математическая статистика [Текст] учеб. пособие для вузов В. Е. Гмурман. - 12-е изд. - М.: Юрайт, 2014. - 478, [1] с. ил.

б) дополнительная литература:

1. Кремер Н. Ш. Теория вероятностей и математическая статистика : учебник для вузов по экон. специальностям / Н. Ш. Кремер. - 3-е изд., перераб. и доп.. - М. : ЮНИТИ-ДАНА, 2009. - 550, [1] с.

в) отечественные и зарубежные журналы по дисциплине, имеющиеся в библиотеке:

Не предусмотрены

г) методические указания для студентов по освоению дисциплины:

1. Заляпин, В. И. Математическая статистика [Текст] учеб. пособие В. И. Заляпин, Е. В. Харитонов ; Юж.-Урал. гос. ун-т, Каф. Мат. анализ ; ЮУрГУ. - Челябинск: Издательство ЮУрГУ, 2008. - 146 с.

из них: учебно-методическое обеспечение самостоятельной работы студента:

Электронная учебно-методическая документация

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Основная литература	Электронно-библиотечная	Каракулев, Ю. А. Руководство к решению задач с применением электронных таблиц EXCEL : учебное

		система издательства Лань	пособие / Ю. А. Каракулев, А. Н. Иванов. — Санкт-Петербург : НИУ — URL: https://e.lanbook.com/book/43545 (дата обращения: 29.11.2021)
2	Дополнительная литература	Электронно-библиотечная система издательства Лань	Буре, В. М. Методы прикладной статистики в R и Excel : учебное пособие / В. М. Буре, Е. М. Парилина, А. А. Седаков. — 3-е изд., стер. — Санкт-Петербург : Лань, 2019. — 152 с. — ISBN 978-5-8114-2229-6. https://e.lanbook.com/book/112057
3	Основная литература	Электронно-библиотечная система издательства Лань	Степанов, П. Е. Планирование эксперимента : учебно-методическое пособие / П. Е. Степанов. — Москва : МИСИС, 2017. — 22 с. — Текст : электронный // Лань : электронно-библиотечная система. https://e.lanbook.com/book/108113

Перечень используемого программного обеспечения:

1. Microsoft-Windows(бессрочно)
2. Microsoft-Office(бессрочно)

Перечень используемых профессиональных баз данных и информационных справочных систем:

1. -База данных ВИНТИ РАН(бессрочно)

8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Лекции	708а (1)	Мультимедийная аудитория: мультимедийный проектор, персональный компьютер, экран.
Лабораторные занятия	707 (1)	Основное оборудование, компьютерная техника, предустановленное программное обеспечение.
Зачет	707 (1)	Основное оборудование, компьютерная техника, предустановленное программное обеспечение.