

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Нижегородский государственный университет
им. Н.И. Лобачевского»**

Институт информационных технологий, математики и механики

(факультет / институт / филиал)

УТВЕРЖДЕНО
решением ученого совета ННГУ
протокол от
« 30 » _ноября_ 2022 г. № _13_

Рабочая программа дисциплины

Прикладные методы регрессионного и дисперсионного анализа

(наименование дисциплины (модуля))

Уровень высшего образования

магистратура

(бакалавриат / магистратура / специалитет)

Направление подготовки / специальность

02.04.02 Фундаментальная информатика и информационные технологии

(указывается код и наименование направления подготовки / специальности)

Направленность образовательной программы

Вероятностное моделирование и анализ данных

(указывается профиль / магистерская программа / специализация)

Форма обучения

очная

(очная / очно-заочная / заочная)

Нижний Новгород

2023 год

1. Место дисциплины в структуре ОПОП

Дисциплина относится к части, формируемой участниками образовательных отношений
Б1.В.05 Прикладные методы регрессионного и дисперсионного анализа

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами

№ варианта	Место дисциплины в учебном плане образовательной программы	Стандартный текст для автоматического заполнения в конструкторе РПД
2	Блок 1. Дисциплины (модули) Часть, формируемая участниками образовательных отношений	Дисциплина Б1.В.05 «Прикладные методы регрессионного и дисперсионного анализа» относится к части ООП направления подготовки 02.04.02 «Фундаментальная информатика и информационные технологии», формируемой участниками образовательных отношений

достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции		Наименование оценочного средства
	Индикатор достижения компетенции* (код, содержание индикатора)	Результаты обучения по дисциплине**	
ПК-11 Способен определять сферы применения результатов научно-исследовательских и опытно-конструкторских работ	ПК-11.1 Знать методы определения сферы применения результатов научно-исследовательских и опытно-конструкторских работ	Знать основные понятия теории регрессии и планирования эксперимента	Собеседование Вопросы к экзамену
	ПК-11.2 Уметь определять сферы применения результатов научно-исследовательских и опытно-конструкторских работ	Уметь применять современные информационные технологии в области регрессионного и дисперсионного анализа при решении практических задач	Практические задания
ПК-12 Способен разрабатывать и совершенствовать статистическую теорию в части математической статистики и вероятностных методов анализа числовой и нечисловой	ПК-12.1 Знать методы разработки и совершенствования статистической теории в части математической статистики и вероятностных методов анализа числовой и нечисловой	Знать основные утверждения теории регрессии и планирования эксперимента	Собеседование

информации	информации		
	ПК-12.2 Уметь разрабатывать и совершенствовать статистическую теорию в части математической статистики и вероятностных методов анализа числовой и нечисловой информации	Уметь выбирать статистическую регрессионную модель для анализа числовой и нечисловой информации, проводить регрессионный и дисперсионный анализ и интерпретировать его результаты	Практические задания

3. Структура и содержание дисциплины

3.1. Трудоемкость дисциплины

	Очная форма обучения
Общая трудоемкость	4 ЗЕТ
Часов по учебному плану	144
в том числе	
аудиторные занятия (контактная работа):	34
- занятия лекционного типа	16
- занятия семинарского типа	16
- занятия лабораторного типа	0
- текущий контроль (КСР)	2
самостоятельная работа	74
Промежуточная аттестация – экзамен	36

3.2. Содержание дисциплины

Наименование и краткое содержание разделов и тем дисциплины	Всего (часы) Очная	В том числе				
		Контактная работа (работа во взаимодействии с преподавателем), часы. Из них				Самостоятельная работа обучающегося, часы Очная
		Занятия лекционного типа Очная	Занятия семинарского типа Очная	Занятия лабораторного типа Очная	Всего Очная	
Тема 1. Современные проблемы исследования зависимостей Корреляционный анализ количественных признаков. Оценивание регрессии по обобщенному МНК. Однофакторный, двухфакторный и многофакторный дисперсионный анализ. Планирование дисперсионного эксперимента. Оценка линейных контрастов. Модели со случайными и неслучайными факторами. Категоризованные данные. Анализ таблиц сопряженности признаков 2×2 и $r \times m$.	50	8	8		16	34

Меры связи Чупрова и Гудмена-Краскала.						
Тема 2. Планирование эксперимента Планирование регрессионных экспериментов. Планы с полностью перекрестной классификацией. Неполноблочные планы. Дисперсионный анализ при большом числе факторов. Планы 2^k . Алгоритм Йетса	56	8	8		16	40
Текущий контроль (КСР)	2	0	0		2	0
Промежуточная аттестация – экзамен	36	0	0		36	0
Итого	144	16	16		70	74

Практические занятия (семинарские занятия) организуются, в том числе в форме практической подготовки, которая предусматривает участие обучающихся в выполнении отдельных элементов работ, связанных с будущей профессиональной деятельностью.

Практическая подготовка предусматривает: выполнение проекта по статистическому анализу реальных данных методом однофакторного анализа.

На проведение практических занятий (семинарских занятий) в форме практической подготовки отводится 2 часа.

Практическая подготовка направлена на формирование и развитие:

- практических навыков в соответствии с профилем ОП: Применение фундаментальных знаний, полученных в области математических и (или) естественных наук.
- компетенций - ПК-12: Способен разрабатывать и совершенствовать статистическую теорию в части математической статистики и вероятностных методов анализа числовой и нечисловой информации.

Текущий контроль успеваемости реализуется в формах опросов на занятиях семинарского типа.

Промежуточная аттестация проходит в традиционных формах (экзамен)

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

Самостоятельная работа заключается в чтении литературы из списка основной литературы, решения домашних заданий и подготовке в промежуточной аттестации.

Контрольные вопросы и задания для проведения текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведены в п. 5.2.

Для обеспечения самостоятельной работы обучающихся используется электронный курс «Современные проблемы прикладной математической статистики» (<https://e-learning.unn.ru/course/view.php?id=6167>), созданный в системе электронного обучения ННГУ – <https://e-learning.unn.ru/>.

5. Фонд оценочных средств для промежуточной аттестации по дисциплине (модулю), включающий:

5.1. Описание шкал оценивания результатов обучения по дисциплине

Уровень сформированности компетенций (индикатора достижения компетенций)	Шкала оценивания сформированности компетенций						
	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
	Не зачтено		Зачтено				
<u>Знания</u>	Отсутствие знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа	Уровень знаний ниже минимальных требований. Имели место грубые ошибки.	Минимально допустимый уровень знаний. Допущено много негрубых ошибок.	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько незначительных ошибок	Уровень знаний в объеме, соответствующем программе подготовки, без ошибок.	Уровень знаний в объеме, превышающем программу подготовки.
<u>Умения</u>	Отсутствие минимальных умений. Невозможность оценить наличие умений вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки.	Продemonстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме.	Продemonстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продemonстрированы все основные умения, решены все основные задачи с отдельными незначительными недочетами, выполнены все задания в полном объеме.	Продemonстрированы все основные умения, решены все основные задачи. Выполнены все задания, в полном объеме без недочетов
<u>Навыки</u>	Отсутствие владения материалом. Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки.	Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами.	Продemonстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продemonстрированы базовые навыки при решении стандартных задач без ошибок и недочетов.	Продemonстрированы навыки при решении нестандартных задач без ошибок и недочетов.	Продemonстрирован творческий подход к решению нестандартных задач.

Шкала оценивания при собеседовании

Результат ответов	Оценка
Студент дает верные развернутые ответы на вопросы	зачтено

преподавателя	
Студент отвечает только на часть вопросов, или отвечает с ошибками, или не дает развернутого ответа на вопросы	не зачтено

Шкала оценки при промежуточной аттестации

Оценка		Уровень подготовки
зачтено	Превосходно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «превосходно»
	Отлично	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично», при этом хотя бы одна компетенция сформирована на уровне «отлично»
	Очень хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо», при этом хотя бы одна компетенция сформирована на уровне «очень хорошо»
	Хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо», при этом хотя бы одна компетенция сформирована на уровне «хорошо»
	Удовлетворительно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
не зачтено	Неудовлетворительно	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»
	Плохо	Хотя бы одна компетенция сформирована на уровне «плохо»

5.2. Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения

5.2.1 Контрольные вопросы

Вопрос	Код компетенции (согласно РПД)
1. Регрессия как условное математическое ожидание. Понятие корреляционной связи. Оптимальное свойство регрессии.	ПК-11
2. Линейная регрессия. Корреляционное отношение Пирсона. Оценка регрессии по методу наименьших квадратов.	ПК-12
3. Однофакторный дисперсионный анализ.	ПК-12
4. Двухфакторный дисперсионный анализ.	ПК-12
5. Анализ таблиц сопряженности признаков. Проверка независимости в таблицах 2×2 . Сравнение медицинских экспериментов.	ПК-12
6. Модели бинарного выбора. Логит- и пробит-модели.	ПК-12

Прогнозирование спроса на рабочую силу	
7. Планирование экспериментов. Планы с полностью перекрестной классификацией. Неполноблочные планы. Планирование с помощью латинских квадратов.	ПК-11
8. Планирование экспериментов. Планы с полностью перекрестной классификацией.	ПК-12
9. Планирование экспериментов. Неполноблочные планы. Планирование с помощью латинских квадратов.	ПК-12
10. Планирование экспериментов. Дисперсионный анализ: факторные планы.. Алгоритм Йетса.	ПК-12

5.2.3. Типовые задания/задачи для оценки сформированности компетенции ПК-11

1. Закон Хаббла в астрономии утверждает, что скорость космологического разбегания галактических скоплений пропорциональна расстоянию до них. Будем считать, что расстояния r (парсек) до скоплений случайны и скорость v (сотни км в сек.) случайна. Рассмотрим данные из *W.L. Freedman et al. Final results from the Hubble space telescope key project to measure the Hubble constant. — Astrophysical Journal, v.553, p.47-72, 2001.*

Таблица 1

№	Список NGC (Созвездие)	r_i (x)	v_i (y)
1.	NGC 0300 (Скульптор)	2.00	133
2.	NGC 0925 (Треугольник)	9.16	664
3.	NGC 1326A (Печь)	16.14	1794
4.	NGC 1365 (Печь)	17.95	1594
5.	NGC 1425 (Печь)	21.88	1473
6.	NGC 2403 (Жираф)	3.22	278
7.	NGC 2541 (Рысь)	11.22	714
8.	NGC 2090 (Голубь)	11.75	882
9.	NGC 3031 (Б.Медведица)	3.63	80
10.	NGC 3198 (Б.Медведица)	13.80	772
11.	NGC 3351 (Лев)	10.00	642
12.	NGC 3368 (Лев)	10.52	768
13.	NGC 3621 (Гидра)	6.64	609
14.	NGC 4321 (Волосы Вероники)	15.21	1433
15.	NGC 4414 (Волосы Вероники)	17.70	619
16.	NGC 4496A	14.86	1424
17.	NGC 4548 (Волосы Вероники)	16.22	1384
18.	NGC 4535 (Дева)	15.78	1444
19.	NGC 4536 (Дева)	14.93	1423
20.	NGC 4639 (Дева)	21.98	1403
21.	NGC 4725 (Волосы Вероники)	12.36	1103

22.	IC 4181 (Волосы Вероники)	4.493.	318
23.	NGC 5253 (Центавр)	3.15	232
24.	NGC 7331 (Пегас)	14.72	999

- 1) Оценить по методу наименьших квадратов зависимость $y = h \cdot x + \varepsilon$;
- 2) Оценить по методу наименьших квадратов зависимость $y = h \cdot x + a + \varepsilon$;
- 3) Проверить гипотезу о степени полинома (в п.1 и п.2) $H_0 : m = 1, H_1 : m \neq 1$.
- 4) По п.1 оценить возраст вселенной, учитывая, что один мегапарсек — это $3.09 \cdot 10^{19}$ км, т.е. в секундах этот возраст равен $t_0 = 3.09 \cdot 10^{19} / \hat{h}$, где \hat{h} — оценка постоянной Хаббла, затем перевести его в млрд.лет, принимая продолжительность года 365 дней.

По современным уточненным данным этот возраст принимается равным 13,798 млрд.лет, а постоянная Хаббла оценивается как 66,93.

5.2.4. Типовые задания/задачи для оценки сформированности компетенции ПК-12

Задача 2. Одной из важных характеристик качества кормовой патоки является число градусов плотности Брикса. Этот показатель количества твердого вещества в патоке и основной фактор, рассматривавшийся при ее производстве. Поставщиками патоки являются три различных района производства. Ниже приведены восемь выборочных показателей для каждого района:

I	II	III
81.6	81.8	82.1
81.3	84.7	79.6
82.0	82.0	83.1
79.6	85.6	80.7
78.4	79.9	81.8
81.8	83.2	79.9
80.2	84.1	82.6
80.7	85.0	81.9

Считая, что $X_{ij}, i = 1, 2, 3; j = 1, \dots, 8$ распределены нормально с одной и той же дисперсией σ^2 проверить гипотезу о том, что данные однородны на уровне значимости $\alpha = 0.05$. 2) Если гипотеза об однородности отвергается, то какой район выделяется из этой совокупности

5.2.5. Типовые вопросы для собеседования для оценки компетенции «ПК-11»

1. Что такое линейная регрессия?
2. Что такое коэффициент корреляции?
3. Что такое план с перекрестной классификацией?

5.2.6. Типовые вопросы для собеседования для оценки компетенции «ПК-12»

1. В чем заключается метод наименьших квадратов?
2. Напишите вид модели однофакторного анализа? Какой вид имеет нулевая гипотеза в этом случае?
3. Напишите вид статистики критерия для проверки независимости в таблице сопряженности

6. Учебно-методическое и информационное обеспечение дисциплины

а) основная литература:

1. Крамер Г. Математические методы статистики. — М.: Мир, 1975. — 648 с. (19 экз.)
2. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. — 2001, М.: ЮНИТИ, т.1 — 656 с. (3 экз.), т.2 — 432 с. (18 экз.)

б) дополнительная литература:

1. Доугерти К. Введение в эконометрику; учебник. —М.: Инфра-М, 2004. — 419 с. (52 экз.)

2. Зорин А.В., Федоткин М.А. Введение в прикладной статистический анализ в пакете R: Учебно-методическое пособие. — Нижний Новгород: ННГУ, 2010. — 50 с.
<http://www.unn.ru/pages/e-library/methodmaterial/files/3.pdf> (Фонд Учебно-методических материалов, подготовленных в рамках реализации проекта "Нижегородский государственный университет им. Лобачевского - Национальный исследовательский университет")
- в) программное обеспечение и Интернет-ресурсы (в соответствии с содержанием дисциплины):
- Сайт проекта R для установки свободного программного продукта <http://www.r-project.org>

7. Материально-техническое обеспечение дисциплины

Помещения представляют собой учебные аудитории для проведения учебных занятий, предусмотренных программой, оснащенные оборудованием и техническими средствами обучения

- Операционные системы семейства Microsoft Windows, лицензия по подписке Microsoft Imagine.
- Любая свободно распространяемая реализация языка R (лицензия GNU).

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечены доступом в электронную информационно-образовательную среду

Программа составлена в соответствии с требованиями ОС ННГУ по направлению 02.04.02 Фундаментальная информатика и информационные технологии.

Автор д.ф.-м.н., профессор _____ М.С.Тихов

Рецензент (ы) _____

Заведующий кафедрой _____ А.В. Зорин

Программа одобрена на заседании методической комиссии института информационных технологий, математики и механики
от 30.11.2022 года, протокол № 3.