

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 многомерные методы статистики

(наименование дисциплины (модуля))

Уровень высшего образования

магистратура

(бакалавриат / магистратура / специалитет)

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.ДВ.02.01 Прикладные многомерные методы статистики

№ 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.ДВ.02.01, «Прикладные многомерные методы статистики» относится к части ООП направления подготовки 02.04.02 «Фундаментальная информатика и информационные технологии», формируемой участниками образовательных отношений

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами 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: знать типовые прикладные задачи многомерного статистического анализа	Собеседование
	ПК-11.2 Уметь определять сферы применения результатов научно-исследовательских и опытно-конструкторских работ	Уметь выполнять типовые процедуры многомерного статистического анализа для научно-исследовательских работ	Задачи
ПК-12 Способен разрабатывать и совершенствовать статистическую теорию в части математической статистики и вероятностных методов анализа числовой и нечисловой информации	ПК-12.1 Знать методы разработки и совершенствования статистической теории в части математической статистики и вероятностных методов анализа числовой и нечисловой информации	Знать алгоритмы решения статистических задач многомерного статистического анализа	Собеседование
	ПК-12.2 Уметь разрабатывать и совершенствовать	Уметь использовать современное программное обеспечение для решения прикладных задач многомерного	Задачи

	статистическую теорию в части математической статистики и вероятностных методов анализа числовой и нечисловой информации	статистического анализа	
--	--	-------------------------	--

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):	33
- занятия лекционного типа	16
- занятия семинарского типа	16
- занятия лабораторного типа	0
- текущий контроль (КСР)	1
самостоятельная работа	111
Промежуточная аттестация – зачет	

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	
1. Распознавание образов и типологизация объектов. Классификация с обучением: дискриминантный анализ. Классификация без обучения: кластерный анализ	47	8	8		16	31
2. Снижение размерности признакового пространства. Факторный анализ. Метод главных компонент.	52	6	6		12	40
3. Многомерное шкалирование. Задача многомерного шкалирования. Меры близости на основе условных вероятностей. Метрические меры различия профилей. Модель Торнгенсона. Неметрические методы. Снижение размерности признакового пространства.	44	2	2		4	40
Текущий контроль (КСР)	1	0	0		1	0
Промежуточная аттестация – зачет						
Итого	144	16	16		33	111

Текущий контроль успеваемости реализуется в формах опросов на занятиях семинарского типа

Промежуточная аттестация проходит в традиционных форма (зачет)

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

Самостоятельная работа состоит в чтении основной литературы и подготовке к 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	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	Продemonстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания, в полном	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме, но некоторые с	Продemonстрированы все основные умения, решены все основные задачи с отдельными несущественными недочетами, выполнены	Продemonстрированы все основные умения, решены все основные задачи. Выполнены все задания, в полном объеме без

	ся от ответа		полном объеме.	объеме, но некоторые с недочетами.	недочетами.	все задания в полном объеме.	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.	Продемонстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продемонстрированы базовые навыки при решении стандартных задач без ошибок и недочетов.	Продемонстрированы навыки при решении нестандартных задач без ошибок и недочетов.	Продемонстрирован творческий подход к решению нестандартных задач.

Шкала оценивания при собеседовании

Результат ответов	Оценка
Студент дает верные развернутые ответы на вопросы преподавателя	зачтено
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. Дискриминантный анализ. Проблемы классификации. Принципы правильной классификации. Функции потерь и вероятности неправильной классификации. Решение задачи байесовских процедур классификации.	ПК-11
2. Статистическая оценка надежности решения методом главных компонент.	ПК-11
3. Факторный анализ. Факторные нагрузки. Простая факторная структура. Вращение факторной структуры. Методы вращения.	ПК-11
4. Факторный анализ – метод главных компонент. Собственные значения, задача о числе факторов. Критерий Кайзера. Критерий каменистой осыпи.	ПК-11
5. Дискриминантный анализ. Определение ошибки классификации.	ПК-11
6. Дискриминантный анализ. Классификация без обучения: расщепление смесей распределения. EM-алгоритм.	ПК-12
7. Дискриминантный анализ. Классификация наблюдений в случае нескольких генеральных совокупностей. Классификация с помощью функции расстояния.	ПК-11
8. Дискриминантный анализ в случае нормальных распределений. Две генеральные совокупности с неравными ковариациями.	ПК-11
9. Дискриминантный анализ в случае нормальных распределений. Две генеральные совокупности с равными ковариациями.	ПК-11
10. Дискриминантный анализ. Принципы правильной классификации. Функции потерь и вероятности неправильной классификации. Решение задачи байесовских процедур классификации.	ПК-12
11. Иерархический кластерный анализ. Дендограммы. Примеры.	ПК-11
12. Кластерный анализ. Функционалы качества разбиений. Число кластеров.	ПК-11
13. Кластерный анализ. Метод средних. Метод поиска сгущений. Метод Ворда.	ПК-11
14. Кластерный анализ. Меры сходства между объектами. Расстояние между кластерами.	ПК-12
15. Применение кластерного и дискриминантного анализа.	ПК-11

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

Задача 1. Деятельность пяти сельскохозяйственных предприятий характеризуется показателями объема реализованной продукции $x^{(1)}$ – растениеводства и $x^{(2)}$ – животноводства с одного гектара пашни (тыс.руб..га). Значения показателей представлены в таблице:

Номер хозяйства (i)	1	2	3	4	5
$x^{(1)}$	24,9	15,1	11,7	16,7	27,3
$x^{(2)}$	9,8	11,1	8,8	8,9	9,4

Требуется с помощью иерархического агломеративного алгоритма провести классификацию сельскохозяйственных предприятий и построить дендограмму при использовании обычной евклидовой метрики методом «дальнего соседа». Сделать выводы.

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

Задача 1. По семи предприятиям имеются следующие данные о результатах работы за отчетный период:

Таблица

Номер предприятия	Выпуск продукции на одного работающего, млн. руб.	Прибыль от реализации продукции млн. руб.
1	51	28
2	63	39
3	48	29
4	39	37
5	30	18
6	58	36
7	61	55

По иерархическому агломеративному алгоритму провести классификацию $n = 7$ предприятий. В качестве расстояния между объектами принять евклидово расстояние. Расстояние между кластерами измерять по принципу «средней связи».

5.2.4. Вопросы для собеседования

Вопросы для оценки компетенции «ПК-11»

1. Многомерный статистический анализ: его особенности и технологии. Основные разделы многомерного статистического анализа.
2. Кластерный анализ. Метод k – средних. Метод поиска сгущений. Функционалы качества разбиения. Число кластеров.
3. Кластерный анализ. Иерархический кластерный анализ. Дендограммы.
4. Примеры применения кластерного анализа. Использование кластерного анализа для характеристики деятельности коммерческих банков на основе данных журнала "Профиль".
5. Дискриминантный анализ. Параметрический дискриминантный анализ в случае нормальных классов Две генеральные совокупности, имеющие известные многомерные нормальные распределения и равные матрицы ковариаций.
6. Дискриминантный анализ. Параметрический дискриминантный анализ в случае нормальных классов Две генеральные совокупности, имеющие известные многомерные нормальные распределения и неравные матрицы ковариаций.
7. Дискриминантный анализ. Параметрический дискриминантный анализ в случае нормальных классов Многомерные нормальные распределения, параметры которых оцениваются по выборке.

8. Дискриминантный анализ. Классификация наблюдений в случае нескольких генеральных совокупностей. Определение числа и вида дискриминирующих функций. Классификация объектов с помощью функции расстояния.
9. Применение дискриминантного анализа: диагностика причин кризисного состояния и банкротства предприятий.
10. Факторный анализ. Основные понятия факторного анализа.
11. Метод главных компонент. Статистическая оценка надежности решений методами главных компонент.
12. Задача о числе факторов. Критерий Кайзера. Критерий «каменистой осыпи».
13. Факторный анализ как метод классификации. Факторные нагрузки. Вращение факторной структуры. Методы вращения.
14. Многомерное шкалирование. Задача многомерного шкалирования. Основные подходы к многомерному шкалированию. Меры близости на основе условных вероятностей.
15. Многомерное шкалирование. Метрические меры различия профилей. Модель Торнгенсона. Примеры.
16. Неметрическое многомерное шкалирование.
17. Использование многомерного шкалирования в маркетинговых исследованиях

6. Учебно-методическое и информационное обеспечение дисциплины

а) основная литература:

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика. Основы эконометрики: учеб. для вузов: в 2 т.. – М.: ЮНИТИ, 2001. т.1 – 656 с (3 экз), т.2 – 432 с. (18 экз).
2. Дубров А.М., Мхитарян В.С., Л.И. Трошин. Многомерные статистические методы. - М., Финансы и статистика, 2000. – 352 с. (4 экз.)

б) дополнительная литература:

3. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика: начальный курс. - М.: Дело, 2005. – 504 с. (2 экз.).
4. Сошникова Л.А., Тамашевич В.Н., Уеббе Г., Шефер М. Многомерный статистический анализ в экономике. - 1999, М., Юнити, – 598 с. (3 экз).

в) программное обеспечение и Интернет-ресурсы (в соответствии с содержанием дисциплины):

- среда статистических вычислений R (адрес для установки www.r-project.org, лицензия GNU GPL)

7. Материально-техническое обеспечение дисциплины

Помещения представляют собой учебные аудитории для проведения учебных занятий, предусмотренных программой, оснащенные оборудованием и техническими средствами обучения: компьютерный класс, проектор, экран

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.