

Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное учреждение высшего образования
«Новгородский государственный университет имени Ярослава Мудрого»
Институт электронных и информационных систем

Кафедра прикладной математики и информатики



И. Эминов
2019 г.

РАБОЧАЯ ПРОГРАММА
учебной дисциплины
Многомерные статистические методы

по направлению подготовки (специальности)
01.04.02 Прикладная математика и информатика
Направленность (профиль)
Прикладной анализ данных

СОГЛАСОВАНО
Начальник отдела обеспечения
Деятельности ИЭИС
 П.В. Лысухо

«06» апреля 2019 г.

Разработал
профессор КПМИ _____

 В.А. Едемский

«12» февраля 2019 г.

Принято на заседании кафедры
Протокол № 7 от «24» 02 2019 г.
Заведующий кафедрой

 А. В. Колногоров
«27» февраля 2019 г.

1 Цели и задачи освоения учебной дисциплины

Цель освоения учебной дисциплины: формирование компетентности студентов в области совершенствования и реализации новых математических методов решения прикладных задач; разработки математических моделей и их анализа при решении задач в области профессиональной деятельности.

Задачи:

- a) формирование у студентов необходимого объема знаний по основам многомерного статистического анализа;
- b) ознакомление студентов с особенностями факторного, кластерного и дискриминантного анализов;
- c) овладение умениями решать практические задачи и производить обработку данных на ЭВМ;
- d) выработка навыков активного применения пакетов прикладных статистических программ для ПЭВМ.

2 Место учебной дисциплины в структуре ОПОП

Учебная дисциплина относится к обязательной части учебного плана основной профессиональной образовательной программы направления подготовки (специальности) 01.04.02 Прикладная математика и информатика и направленности (профилю) Прикладной анализ данных (далее – ОПОП). В качестве входных требований выступают сформированные ранее компетенции обучающихся, приобретенные ими в рамках следующих дисциплин (модулей, практик): Алгебра, геометрия и математическая логика, Теория вероятностей и математическая статистика. Освоение учебной дисциплины является компетентностным ресурсом для дальнейшего изучения следующих дисциплин (модулей, практик): Теория и моделирование систем массового обслуживания, Прикладной стохастический анализ, Научно-исследовательская работа.

3 Требования к результатам освоения учебной дисциплины

Перечень компетенций, которые формируются в процессе освоения учебной дисциплины:

Общепрофессиональные компетенции:

ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач

ОПК-3 Способен разрабатывать математические модели и проводить их анализ при решении задач в области профессиональной деятельности

Таблица 1 - Результаты освоения учебной дисциплины

Код и наименование компетенции	Результаты освоения учебной дисциплины (индикаторы достижения компетенций)		
ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	Знает современные математические методы решения прикладных задач	Умеет совершенствовать и реализовывать новые математические методы при решении задач в профессиональной	Владеет навыками выбора и применения математических методов решения задач прикладного анализа данных

		области	
ОПК-3 Способен разрабатывать математические модели и проводить их анализ при решении задач в области профессиональной деятельности	Знает математические и имитационные модели, используемые при решении задач в области профессиональной деятельности	Умеет разрабатывать и применять математические и имитационные модели при решении задач прикладного анализа данных	Владеет навыками математического и имитационного моделирования

4 Структура и содержание учебной дисциплины

4.1 Трудоемкость учебной дисциплины

Таблица 2 - Трудоемкость учебной дисциплины для очной формы обучения

Части учебной дисциплины	Всего	Распределение по семестрам
		3 семестр
1. Трудоемкость учебной дисциплины в зачетных единицах (ЗЕТ)	6	6
2. Контактная аудиторная работа в академических часах (АЧ)	45	45
3. Курсовая работа/курсовой проект (АЧ) <i>(при наличии)</i>		
4. Внеаудиторная СРС в академических часах (АЧ)	135	135
5. Промежуточная аттестация экзамен (АЧ)	36	36

4.2 Содержание учебной дисциплины

Раздел № 1 Статистическое оценивание, сравнение многомерных генеральных совокупностей и компонентный анализ

1.1. Точечные оценки параметров многомерной генеральной совокупности. Доверительные области. Проверка гипотез. Проверка гипотез о равенстве вектора средних значений постоянному вектору. Проверка гипотез о равенстве двух векторов средних значений.

1.2. Статистический подход в методе главных компонент. Многомерное нормальное распределение как модель. Линейная модель метода главных компонент. Получение матрицы весовых коэффициентов по алгоритму метода главных компонент. Дисперсия исследуемых признаков в методе главных компонент.

Раздел № 2 Факторный анализ

2.1. Сущность методов факторного анализа и их классификация, основная модель факторного анализа. Компоненты дисперсии в факторном анализе. Получение матрицы коэффициентов парной корреляции в факторном анализе.

2.2. Метод главных факторов и его алгоритм. Метод максимального правдоподобия. Вращение пространства общих факторов, проблемы вращения.

2.3. Факторный анализ и методы классификации многомерных наблюдений. Статистическая оценка надежности решений методами главных компонент и факторного анализа.

Раздел № 3 Кластерный анализ

3.1. Классификация без обучения. Иерархические кластер - процедуры.

3.2. Метод k - средних. Функционалы качества разбиения.

Раздел № 4 Дискриминантный анализ

4.1. Методы классификации с обучением. Линейный дискриминантный анализ.

4.2. Дискриминантный анализ при нормальном законе распределения показателей.

4.3 Трудоемкость разделов учебной дисциплины и контактной работы

Таблица 3 - Трудоемкость разделов учебной дисциплины

№	Наименование разделов (тем) учебной дисциплины (модуля), УЭМ, наличие КП/КР	Контактная работа (в АЧ)				Внеауд. СРС (в АЧ)	Формы текущего контроля
		Аудиторная			В т.ч. СРС		
		ЛЕК	ПЗ	ЛР			
1.	Статистическое оценивание, сравнение многомерных генеральных совокупностей и компонентный анализ	2	9		2	30	ИДЗ№1 КР №1
2.	Факторный анализ	3	9		3	45	ИДЗ№2 КР №2
3.	Кластерный анализ	2	9		2	30	ИДЗ№3 КР №3
4.	Дискриминантный анализ	2	9		2	30	ИДЗ№4 КР №4
	Промежуточная аттестация					36	экзамен
	ИТОГО	9	36		9	135	

4.4 Лабораторные работы и курсовые работы/курсовые проекты

4.4.1 Перечень тем лабораторных работ:

Лабораторные работы не предусмотрены учебным планом.

4.4.2 Примерные темы курсовых работ/курсовых проектов:

Курсовые работы/курсовые проекты не предусмотрены учебным планом.

5 Методические рекомендации по организации освоения учебной дисциплины

Таблица 4 - Методические рекомендации по организации лекций

№	Темы лекционных занятий (форма проведения)	Трудоемкость в АЧ
1.	Статистическое оценивание, сравнение многомерных генеральных совокупностей и компонентный анализ (информационная лекция)	2
2.	Факторный анализ (лекция-презентация)	3
3.	Кластерный анализ (информационная лекция)	2
4.	Дискриминантный анализ (лекция-презентация)	2
	ИТОГО	9

Таблица 5 - Методические рекомендации по организации практических занятий

№	Темы практических занятий (форма проведения)	Трудоемкость в АЧ
1.	Точечные оценки параметров многомерной генеральной совокупности. Доверительные области (решение задач. Работа в группах. Формирование умений и навыков решения задач по теме).	3
2.	Проверка гипотез. Проверка гипотез о равенстве вектора средних значений постоянному вектору. Проверка гипотез о равенстве двух векторов средних значений (решение задач. Работа в группах. Формирование умений и навыков решения задач по теме).	3
3.	Статистический подход в методе главных компонент. Линейная модель метода главных компонент. Получение матрицы весовых коэффициентов по алгоритму метода главных компонент. Дисперсия исследуемых признаков в методе главных компонент (решение задач. Работа в группах. Формирование умений и навыков решения задач по теме).	3
4.	Сущность методов факторного анализа и их классификация, основная модель факторного анализа (решение задач. Работа в группах. Формирование умений и навыков решения задач по теме).	3
5.	Компоненты дисперсии в факторном анализе. Получение матрицы коэффициентов парной корреляции в факторном анализе (решение задач. Работа в группах. Формирование умений и навыков решения задач по теме).	3
6.	Метод максимального правдоподобия. Вращение пространства общих факторов, проблемы вращения. Факторный анализ и методы классификации многомерных наблюдений. Статистическая оценка надежности решений методами главных компонент и факторного анализа (решение задач. Работа в группах. Формирование умений и навыков решения задач по теме).	3
7.	Классификация без обучения (решение задач. Работа в группах. Формирование умений и навыков решения задач по теме).	3
8.	Иерархические кластер – процедуры (решение задач. Работа в группах. Формирование умений и навыков решения задач по теме).	3
9.	Метод k - средних. Функционалы качества разбиения (решение задач. Работа в группах. Формирование умений и навыков решения задач по теме).	3
10.	Методы классификации с обучением (решение задач. Работа в группах. Формирование умений и навыков решения задач по теме).	3
11.	Линейный дискриминантный анализ (решение задач. Работа в группах. Формирование умений и навыков решения задач по теме).	3
12.	Дискриминантный анализ при нормальном законе распределения показателей (решение задач. Работа в группах. Формирование умений и навыков решения задач по теме).	3
	ИТОГО	36

6 Фонд оценочных средств учебной дисциплины

Фонд оценочных средств представлен в Приложении А.

7 Условия освоения учебной дисциплины

7.1 Учебно-методическое обеспечение

Учебно-методического обеспечение учебной дисциплины представлено в Приложении Б.

7.2 Материально-техническое обеспечение

Таблица 6 - Материально-техническое обеспечение учебной дисциплины

№	Требование к материально-техническому обеспечению	Наличие материально-технического оборудования и программного обеспечения
1.	Наличие учебной аудитории	Учебная мебель, доска
2.	Мультимедийное оборудование	1 компьютер, проектор, экран, выход в интернет
3.	Программное обеспечение	Microsoft Windows XP Professional. Лицензия «Open License» № 45257130; Microsoft Office 2007. Лицензия «Open License» № 47742190.

Приложение А
(обязательное)
Фонд оценочных средств
учебной дисциплины многомерные статистические методы

1 Структура фонда оценочных средств

Фонд оценочных средств состоит из двух частей:

а) открытая часть - общая информация об оценочных средствах (название оценочных средств, проверяемые компетенции, баллы, количество вариантов заданий, методические рекомендации для применения оценочных средств и пр.), которая представлена в данном документе, а также те вопросы и задания, которые могут быть доступны для обучающегося;

б) закрытая часть - фонд вопросов и заданий, которая не может быть заранее доступна для обучающихся (экзаменационные билеты, вопросы к контрольной работе и пр.) и которая хранится на кафедре.

2 Перечень оценочных средств текущего контроля и форм промежуточной аттестации

Таблица А.1 - Перечень оценочных средств

№	Оценочные средства для текущего контроля	Разделы (темы) учебной дисциплины	Баллы	Проверяемые компетенции
1.	Индивидуальное домашнее задание № 1	Статистическое оценивание, сравнение многомерных генеральных совокупностей и компонентный анализ	20	ОПК-2, ОПК-3
2.	Контрольная работа №1	Статистическое оценивание, сравнение многомерных генеральных совокупностей и компонентный анализ	40	ОПК-2, ОПК-3
3.	Индивидуальное домашнее задание № 2	Факторный анализ	25	ОПК-2, ОПК-3
4.	Контрольная работа №2	Факторный анализ	45	ОПК-2, ОПК-3
5.	Индивидуальное домашнее задание № 3	Кластерный анализ	20	ОПК-2, ОПК-3
6.	Контрольная работа №3	Кластерный анализ	40	ОПК-2, ОПК-3
7.	Индивидуальное домашнее задание № 4	Дискриминантный анализ	20	ОПК-2, ОПК-3
8.	Контрольная работа №4	Дискриминантный анализ	40	ОПК-2, ОПК-3
<i>Промежуточная аттестация</i>				
	Экзамен		50	
	ИТОГО		300	

3 Рекомендации к использованию оценочных средств

1) Индивидуальные домашние задания

Таблица А.2 – Индивидуальные домашние задания

<i>Критерии оценки</i>	<i>Количество вариантов заданий</i>	<i>Количество вопросов</i>
Количество правильных ответов	25	4-5
Использование терминологии		
Обоснованность решений		

Демонстрационные варианты индивидуальных домашних заданий

ИДЗ №1 по теме “Статистическое оценивание, сравнение многомерных генеральных совокупностей и компонентный анализ”.

1. Найти оценки математических ожиданий и корреляционной матрицы системы четырех случайных величин;

2. Найти эллипсоидную и прямоугольную доверительные области для вектора математических ожиданий первых двух случайных величин с надежностью $\gamma = 0.95$;

3. Для четырех случайных величин найти компонентную матрицу и вклад в дисперсию каждой компоненты.

ИДЗ №2 по теме “Факторный анализ”.

1. Провести факторный анализ, не совершая поворот факторов. Получить матрицу парных коэффициентов корреляции. Выделить общие факторы, дающие вклад в суммарную дисперсию 90%. Дать экономическую интерпретацию полученных результатов.

2. На основе решения полученного в задании 1 произвести поворот факторов различными методами. Выполнить сравнительный анализ полученных результатов.

3. Оценить адекватность полученной модели, применяя критерий Барлетта.

ИДЗ №3 по теме “Кластерный анализ”.

1. Провести классификацию объектов иерархическим методом. В качестве расстояния между объектами принять “обычное евклидово расстояние”, а расстояния между кластерами измерять по методам: «ближайшего соседа», «дальнего соседа», «средних связей».

2. Решить задание 1 методом Уорда.

3. Решить задание 1 методом К-средних.

4. Построить дендограммы.

ИДЗ №4 по теме “Дискриминантный анализ”.

1. Разбить 20 объектов на три группы, применяя один из методов кластерного анализа. Найти линейную дискриминантную функцию, разделяющую полученные группы объектов.

2. Провести классификацию оставшихся объектов.

3. Найти коэффициенты канонической корреляции между двумя группами.

2) Контрольные работы

Таблица А.3 – Контрольные работы

<i>Критерии оценки</i>	<i>Количество вариантов заданий</i>	<i>Количество вопросов</i>
Количество правильных ответов	25	4-5
Использование терминологии		
Обоснованность решений		
Демонстрация знания основных понятий раздела		

Демонстрационные варианты контрольных работ

Контрольная работа по теме “Статистическое оценивание, сравнение многомерных генеральных совокупностей и компонентный анализ”.

1. Найти оценки математических ожиданий и корреляционной матрицы системы случайных величин X_1, X_2 .

2. Найти доверительную область для вектора математических ожиданий X_1, X_2 с надежностью $\gamma = 0.95$.

3. Для X_1, X_2 найти компонентную матрицу и вклад в дисперсию каждой компоненты. Изобразить новую систему координат на плоскости.

Контрольная работа по теме “Факторный анализ”.

1. По исходным данным провести факторный анализ, не совершая поворот факторов. Получить матрицу парных коэффициентов корреляции. Получить первые три общие фактора.

2. На основе решения полученного в Задании 1 произвести поворот факторов методом Варимакс нормализованный.

3. Оценить адекватность полученной модели, применяя критерий Бартлетта.

Контрольная работа по теме “Кластерный анализ”.

1. Провести классификацию объектов иерархическим методом. В качестве расстояния между объектами принять “обычное евклидово расстояние”, а расстояния между кластерами измерять по принципу: “ближайшего соседа”.

2. Решить задание 1 при условии, что расстояния между кластерами измеряются по принципу “дальнего соседа”.

3. Решить задание 1 методом К-средних.

4. Построить дендограммы.

Контрольная работа по теме “Дискриминантный анализ”.

1. Найти линейные дискриминантные функции, разделяющие заданные группы объектов.

2. Провести классификацию указанных объектов.

3. Найти коэффициент канонической корреляции между двумя группами.

Приведенные примеры позволяют студентам оценить степень сложность заданий, которые им предстоит выполнить на практическом занятии и во время контрольных работ.

3) Экзамен

Таблица А.4 – Экзамен

<i>Критерии оценки</i>	<i>Количество вариантов заданий</i>	<i>Количество вопросов</i>
Количество правильных ответов	25	4-5
Использование терминологии		
Владение материалом		
Логичность и обоснованность ответа		

Список контрольных вопросов к экзамену

1. Оценки параметров распределения.
2. Доверительные области для параметров нормального распределения.
3. Проверка гипотез о выборках.
4. Проверка гипотез о параметрах нормального распределения.
5. Выделение главной компоненты.
6. Метод главных компонент, свойства компонентной матрицы.
7. Общая модель факторного анализа. Метод главных факторов.
8. Факторный анализ и уравнение регрессии.
9. Иерархические кластерные методы.
10. Метод k-средних. Метод поиска сгущений.
11. Дискриминантный анализ.
12. Линейный дискриминант Фишера.
13. Метод канонических корреляций.

Пример экзаменационного билета:

Новгородский государственный университет имени Ярослава Мудрого

Кафедра прикладной математики и информатики

Экзаменационный билет № __ 1 __

Учебная дисциплина многомерные статистические методы

Для направления подготовки (специальности) 01.04.02 Прикладная математика и информатика

1. Доверительные области для параметров нормального распределения.
2. Линейный дискриминант Фишера.
3. Задача.

Принято на заседании кафедры « _____ » _____ 20__ г. Протокол № _____
Заведующий кафедрой _____ (Колногоров А.В.)

Все материалы для проведения промежуточного контроля хранятся на кафедре.

Приложение Б
(обязательное)

Карта учебно-методического обеспечения
Учебной дисциплины многомерные статистические методы

1. Основная литература*

Библиографическое описание издания (автор, наименование, вид, место и год издания, кол. стр.)	Кол. экз. в библ. НовГУ	Наличие в ЭБС
Печатные источники		
1. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы для экономистов и менеджеров : учеб. для студентов экон. спец. вузов. - М. : Финансы и статистика, 2000. - 350с.	23	Национальная электронная библиотека (НЭБ)
2		
Электронные ресурсы		
Электронная библиотека механико-математического факультета Московского государственного университета [Электронный ресурс] : [официальный сайт] / Московск. гос. ун-т. –М., 2002-2015. –Режим доступа: http://www.http://lib.mexmat.ru/ , свободный. -Загл. с экрана (дата обращения: 17.01.2017).	http://poisk.knig.ru	электронная библиотека учебников Мех-Мат МГУ, Москва
Общероссийский математический портал Math-Net.Ru [Электронный ресурс] / Математич. институт им.В.А.СтекловаРАН. -М.; 2002-2015. -Режим доступа: http://http://www.mathnet.ru/about.phtml?option_lang=rus , свободный - Загл. с экрана (дата обращения: 17.01.2017).	http://www.mathnet.ru	общероссийский математический портал

**См. требования п. 4.3.3 ФГОС 3++ (как правило, при использовании в образовательном процессе печатных изданий библиотечный фонд должен быть укомплектован печатными изданиями из расчета не менее 0,25 экземпляра на каждого из изданий, указанных в рабочих программах дисциплин (модулей), на одного обучающегося из числа лиц, одновременно осваивающих соответствующую дисциплину (модуль)).*

2. Дополнительная литература

Библиографическое описание издания (автор, наименование, вид, место и год издания, кол. стр.)	Кол. экз. в библ. НовГУ	Наличие в ЭБС
Печатные источники		
1. Вероятностные разделы математики / Под ред. Ю.Д.Максимова. - СПб.: «Иван Федоров», 2001, 588 с.	13	

Электронные ресурсы		
Интернет-университет информационных технологий [Электронный ресурс] : [официальный сайт] / НОУ «ИНТУ-ИТ», 2003 – 2015. -Режим доступа: http://www.intuit.ru/свободный . -Загл. с экрана (дата обращения: 17.01.2017).	http://www.intuit.ru/	Интернет-университет информационных технологий [Электронный ресурс].

Зав. кафедрой



А.В. Кощов

И.О. Фамилия

« 27 » февраля

2019 г.

