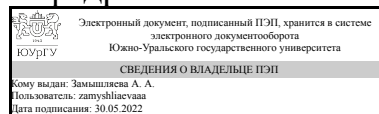


УТВЕРЖДАЮ:  
Заведующий выпускающей  
кафедрой



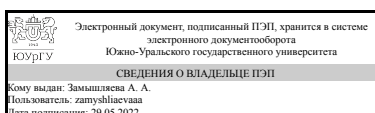
А. А. Замышляева

## РАБОЧАЯ ПРОГРАММА

**дисциплины 1.Ф.М1.07 Статистика для анализа данных  
для направления 09.04.04 Программная инженерия  
уровень Магистратура  
магистерская программа Искусственный интеллект и машинное обучение в  
финтех индустрии  
форма обучения очная  
кафедра-разработчик Прикладная математика и программирование**

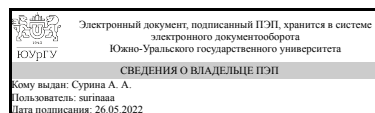
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 09.04.04 Программная инженерия, утверждённым приказом Минобрнауки от 19.09.2017 № 932

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



А. А. Замышляева

Разработчик программы,  
старший преподаватель



А. А. Сурина

## 1. Цели и задачи дисциплины

Целью освоения дисциплины является закрепление и углубление теоретической подготовки обучающегося, развитие у студентов способности к самостоятельным теоретическим и практическим суждениям и выводам, умений объективной оценки научной информации, свободы научного поиска и стремления к применению научных знаний в профессиональной деятельности. Задачи: обеспечения становления профессионального научно-исследовательского мышления студентов, формирование у них четкого представления об основных профессиональных задачах, способах их решения; формирования умений использовать современные технологии сбора информации, обработки и интерпретации полученных экспериментальных и эмпирических данных, владение современными методами исследований; обеспечения готовности к профессиональному самосовершенствованию, развитию инновационного мышления и творческого потенциала, профессионального мастерства; самостоятельного формулирования и решения задач, возникающих в ходе практической деятельности и требующих углубленных профессиональных знаний. В результате освоения дисциплины студент должен получить необходимые сведения для решения следующих профессиональных задач: использование базовых математических задач и математических методов в научных исследованиях; контекстная обработка общенаучной и научно-технической информации, приведение ее к проблемно-задачной форме, анализ и синтез информации.

## Краткое содержание дисциплины

Содержание и назначение многомерного статистического анализа. Многомерный признак и способы его задания. Методы снижения размерности. Статистические методы классификации. Кластерный анализ. Статистические методы классификации. Дискриминантный анализ.

## 2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-5 Способен разрабатывать и применять алгоритмы анализа данных для решения прикладных задач	Знает: назначение, содержание и основные этапы многомерного статистического анализа данных Имеет практический опыт: построения многомерных статистических моделей исследуемых процессов

## 3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
Информационный поиск, анализ и предобработка данных	Математические модели принятия решений в условиях неопределенности, Технологическое предпринимательство

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

Дисциплина	Требования
Информационный поиск, анализ и предобработка данных	Знает: основные принципы сбора, хранения и предобработки данных Умеет: выбирать методы и средства для анализа данных, оценивать возможности и ограничения используемых методов, осуществлять дискретизацию непрерывных данных с учётом решаемой задачи Имеет практический опыт: сбора первичной информации, организации и хранения данных для конкретного исследования, применения методов предобработки данных

#### 4. Объём и виды учебной работы

Общая трудоёмкость дисциплины составляет 3 з.е., 108 ч., 40,5 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах	
		Номер семестра	
		2	
Общая трудоёмкость дисциплины	108	108	
<i>Аудиторные занятия:</i>	32	32	
Лекции (Л)	16	16	
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	16	16	
Лабораторные работы (ЛР)	0	0	
<i>Самостоятельная работа (СРС)</i>	67,5	67,5	
Подготовка отчетов по лабораторным работам	19,5	19,5	
Проработка лекционного материала. Подготовка к выполнению лабораторных работ.	24	24	
Подготовка к экзамену	24	24	
Консультации и промежуточная аттестация	8,5	8,5	
Вид контроля (зачет, диф.зачет, экзамен)	-	экзамен	

#### 5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Содержание и назначение многомерного статистического анализа	2	2	0	0
2	Многомерный признак и способы его задания	8	4	4	0
3	Методы снижения размерности	6	2	4	0
4	Статистические методы классификации. Кластерный анализ	10	6	4	0

5	Статистические методы классификации. Дискриминантный анализ	6	2	4	0
---	--	---	---	---	---

## 5.1. Лекции

№ лекции	№ раздела	Наименование или краткое содержание лекционного занятия	Кол-во часов
1	1	Содержание и назначение многомерного статистического анализа	2
2	2	Многомерные (совместные), частные (маргинальные) и условные законы распределения вероятностей многомерного признака.	2
3	2	Многомерное нормальное распределение и его свойства. Оценки параметров многомерного нормального распределения. Измерители степени тесноты статистической связи между компонентами многомерного признака (корреляционный анализ многомерной выборки): парные, частные и множественные коэффициенты корреляции, корреляционное отношение; ранговые корреляции; анализ таблиц сопряженности и информационная мера связи.	2
4	3	Общая экстремальная постановка задачи снижения размерности и два подхода к определению критерия информативности показателей. Основные методы снижения размерности, нацеленные на оптимизацию критериев автоинформативности: метод главных компонент, факторный анализ, метод экстремальной группировки признаков, методы целенаправленного проецирования исходных многомерных данных.	2
5	4	Общая экстремальная постановка задачи классификации. Две основные формы задания исходной информации в задачах классификации. Основные типы задач классификации.	2
6	4	Основные типы расстояний между объектами и между классами объектов, используемые в процедурах классификации. Обобщенные расстояния Колмогорова.	2
7	4	Параметрические и непараметрические методы классификации без обучающих выборок: оценка параметров смеси распределений, методы кластер-анализа (в том числе — метод k-средних и иерархические кластер-процедуры). Задачи типологизации объектов.	2
8	5	Параметрические и непараметрические методы классификации при наличии обучающих выборок (методы дискриминантного анализа). Задачи социально-экономической диагностики. Связь дискриминантного анализа с логит- и пробит- моделями.	2

## 5.2. Практические занятия, семинары

№ занятия	№ раздела	Наименование или краткое содержание практического занятия, семинара	Кол-во часов
1	2	Проведение экономико-статистических расчетов в пакете анализа MS Excel: статистические функции, ковариационный анализ, корреляционный анализ	2
2	2	Проведение множественного регрессионного анализа в пакете STATISTICA для изучения взаимосвязи переменных экономической задачи	2
3, 4	3	Проведение множественного регрессионного анализа в пакете STATISTICA для определения степени важности изучаемых величин и уменьшения размерности фазового пространства экономической задачи	4
5	4	Кластерный анализ и его реализация в пакете STATISTICA. Иерархические алгоритмы	2

6	4	Кластерный анализ и его реализация в пакете STATISTICA. Процедуры эталонного типа	2
7, 8	5	Дискриминантный анализ и его реализация в пакете STATISTICA	4

### 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 отчетов по лабораторным работам	Шашок А.В. Лабораторный практикум в системе «STATISTICA»: Методические указания к выполнению лабораторных работ по дисциплине «Статистика» для студентов всех форм обучения специальности 080507 «Менеджмент организации» / Рубцовский индустриальный институт. – Рубцовск, 2010. – 81 с.	2	19,5
Проработка лекционного материала. Подготовка к выполнению лабораторных работ.	1. Козлов А.Ю. Статистический анализ данных в MS Excel: Учебное пособие / Козлов А.Ю., Мхитарян В.С., Шишов В.Ф. - М.:НИЦ ИНФРА-М, 2016. - 320 с. 2. Горицкий, Ю. А. Практикум по статистике с пакетами STATGRANICS, STATISTICA, SPSS: Учебное пособие по курсу "Математическая статистика" / Ю. А. Горицкий, Е. Е. Перцов. – М. : Изд-во МЭИ, 1997 . – 84 с.	2	24
Подготовка к экзамену	1. Лемешко Б.Ю. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов и др. - М.: НИЦ ИНФРА-М, 2015. - 890 с. 2. Гирко, В. Л. Многомерный статистический анализ Учеб. пособие для вузов по спец."Математика" и "Прикладная математика ". - Киев: Выща школа, 1988. - 319 с. 3. Андерсон, Т. Введение в многомерный статистический анализ Пер. с англ.: Ю. Ф. Кичатова и др.; Под ред. Б. В. Гнеденко. - М.: Физматгиз, 1963. - 500 с. 4. Дубровский, С. А. Прикладной многомерный статистический анализ. - М.: Финансы и статистика, 1982. - 216 с.	2	24

### 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	2	Текущий контроль	ЛР 1. Статистические методы обработки экспериментальных данных.	1	5	Использованы основные методы статистической обработки данных - 2 балла Проведена оценка параметров реализуемых моделей и методов - 1 балл Выполнен анализ результатов, выводы логичны, использованы стандартные программные средства - 1 балл Правильные ответы на вопросы – 1 балл Иначе - 0 баллов	экзамен
2	2	Текущий контроль	ЛР 2. Проведение экономико-статистических расчетов в пакете анализа MS Excel: многофакторный дисперсионный анализ	1	5	Использованы основные методы статистической обработки данных - 2 балла Проведена оценка параметров реализуемых моделей и методов - 1 балл Выполнен анализ результатов, выводы логичны, использованы стандартные программные средства - 1 балл Правильные ответы на вопросы – 1 балл Иначе - 0 баллов	экзамен
3	2	Текущий контроль	ЛР 3. Парный корреляционный анализ	1	10	Использованы основные методы статистической обработки данных - 3 балла Проведена оценка параметров реализуемых моделей и методов - 3 балла Выполнен анализ результатов, выводы логичны, использованы стандартные программные средства - 2 балла Правильные ответы на вопросы – 2 балла Иначе - 0 баллов	экзамен
4	2	Текущий контроль	ЛР 4. Решение задачи линейного корреляционного и регрессионного анализа	1	5	Использованы основные методы статистической обработки данных - 2 балла Проведена оценка параметров реализуемых моделей и методов - 1 балл	экзамен

						Выполнен анализ результатов, выводы логичны, использованы стандартные программные средства - 1 балл Правильные ответы на вопросы – 1 балл Иначе - 0 баллов	
5	2	Текущий контроль	ЛР 5. Оценивание многомерных линейных регрессионных моделей методами наименьших квадратов и наименьших модулей	1	5	Использованы основные методы статистической обработки данных - 2 балла Проведена оценка параметров реализуемых моделей и методов - 1 балла Выполнен анализ результатов, выводы логичны, использованы стандартные программные средства - 1 балл Правильные ответы на вопросы – 1 балл Иначе - 0 баллов	экзамен
6	2	Текущий контроль	ЛР 6. Факторный анализ	1	5	Использованы основные методы статистической обработки данных - 5 балла Проведена оценка параметров реализуемых моделей и методов - 1 балл Выполнен анализ результатов, выводы логичны, использованы стандартные программные средства - 1 балл Правильные ответы на вопросы – 1 балл Иначе - 0 баллов	экзамен
7	2	Текущий контроль	ЛР 7. Метод канонических корреляций	1	10	Использованы основные методы статистической обработки данных - 3 балла Проведена оценка параметров реализуемых моделей и методов - 3 балла Выполнен анализ результатов, выводы логичны, использованы стандартные программные средства - 2 балла Правильные ответы на вопросы – 2 балла Иначе - 0 баллов	экзамен
8	2	Текущий контроль	ЛР 8. Кластерный анализ	1	5	Использованы основные методы статистической обработки данных - 2 балла Проведена оценка параметров реализуемых моделей и методов - 1 балл Выполнен анализ результатов, выводы логичны, использованы стандартные программные средства - 1 балл	экзамен

						Правильные ответы на вопросы – 1 балл Иначе - 0 баллов	
9	2	Текущий контроль	ЛР 9. Дискриминантный анализ	1	10	Использованы основные методы статистической обработки данных - 3 балла Проведена оценка параметров реализуемых моделей и методов - 6 балла Выполнен анализ результатов, выводы логичны, использованы стандартные программные средства - 2 балла Правильные ответы на вопросы – 2 балла Иначе - 0 баллов	экзамен
10	2	Промежуточная аттестация	Опрос	-	10	Контрольное мероприятие промежуточной аттестации проводится во время экзамена. Студенту задаются 10 вопросов их разных тем курса, позволяющих оценить сформированность компетенций. Правильный ответ на вопрос соответствует 1 баллу. Неправильный ответ на вопрос соответствует 0 баллов.	экзамен

## 6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
экзамен	<p>На экзамене происходит оценивание учебной деятельности обучающихся по дисциплине на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля. Студент вправе пройти контрольное мероприятие в рамках промежуточной аттестации для улучшения своего рейтинга.</p> <p>Студенту задаются 10 вопросов из разных тем курса, позволяющих оценить сформированность компетенций. На подготовку ответов дается не более 20 минут. Использование вспомогательных материалов при подготовке ответа не допускается. По окончании отведенного времени студент озвучивает ответы.</p>	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
ПК-5	Знает: назначение, содержание и основные этапы многомерного статистического анализа данных	+	+	+	+	+	+	+	+	+	+
ПК-5	Имеет практический опыт: построения многомерных статистических моделей исследуемых процессов	+	+			+	+	+	+	+	+



Типовые контрольные задания по каждому мероприятию находятся в приложениях.

## 7. Учебно-методическое и информационное обеспечение дисциплины

### Печатная учебно-методическая документация

#### а) основная литература:

1. Магнус, Я. Р. Эконометрика: Начальный курс Учеб. для вузов по экон. специальностям Акад. нар. хоз-ва при Правительстве Рос. Федерации; Я. Р. Магнус, П. К. Катышев, А. А. Пересецкий. - 3-е изд., перераб. и доп. - М.: Дело, 2000. - 399 с. ил.

#### б) дополнительная литература:

1. Кендалл, М. Многомерный статистический анализ и временные ряды Пер. с англ. Э. Л. Пресмана, В. И. Ротаря; Под ред. А. Н. Колмогорова, Ю. В. Прохорова. - М.: Наука, 1976. - 736 с. ил.
2. Мудров, В. И. Методы обработки измерений: Квазиправдоподобные оценки. - 2-е изд., перераб. и доп. - М.: Радио и связь, 1983. - 304 с. ил.

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

#### г) методические указания для студентов по освоению дисциплины:

1. Методические указания по организации самостоятельной работы студента

#### из них: учебно-методическое обеспечение самостоятельной работы студента:

1. Методические указания по организации самостоятельной работы студента

### Электронная учебно-методическая документация

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Дополнительная литература	Электронно-библиотечная система Znanium.com	Зарова Е.В. Прикладной многомерный статистический анализ: Презентации для лекций и примеры решений с использованием пакета R: Учебное пособие на английском языке / Зарова Е.В. - М.: НИЦ ИНФРА-М, 2016. - 370 с. <a href="https://znanium.com/catalog/product/557578">https://znanium.com/catalog/product/557578</a>
2	Основная литература	Электронно-библиотечная система Znanium.com	Лемешко Б.Ю. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов и др. - М.: НИЦ ИНФРА-М, 2015. - 890 с. <a href="https://znanium.com/catalog/product/515227">https://znanium.com/catalog/product/515227</a>
3	Методические пособия для самостоятельной работы студента	Электронно-библиотечная система Znanium.com	Козлов А.Ю. Статистический анализ данных в MS Excel: Учебное пособие / Козлов А.Ю., Мхитарян В.С., Шишов В.Ф. - М.: НИЦ ИНФРА-М, 2016. - 320 с. <a href="https://znanium.com/catalog/document?id=378179">https://znanium.com/catalog/document?id=378179</a>

4	Основная литература	Российская государственная библиотека	Симчера, В.М. Методы многомерного анализа статистических данных: учебное пособие / В. М. Симчера. - Москва : Финансы и статистика, 2008. - 395 с. <a href="https://dvs.rsl.ru/01005472231">https://dvs.rsl.ru/01005472231</a>
---	---------------------	---------------------------------------	--

Перечень используемого программного обеспечения:

1. Microsoft-Office(бессрочно)

Перечень используемых профессиональных баз данных и информационных справочных систем:

Нет

## 8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Лабораторные занятия	340 (3б)	ПК, MS Office, Foxit Reader (или Adobe Reader)
Лекции	336 (3б)	ПК, проектор, MS Office, Foxit Reader (или Adobe Reader)