

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

ПЕНЗЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ

УТВЕРЖДАЮ
Декан ФВТ

_____ **Фионова Л.Р.**

«_____» _____ 20__ г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Б.1.О.28 Прикладная статистика и интеллектуальный анализ данных

Направление подготовки (специальность) – *09.03.03 Прикладная информатика*

Направленность (профиль подготовки) – *Прикладная информатика в экономике*

Квалификация выпускника – *бакалавр*

Форма обучения – *заочная*

Пенза, 2019

1. Цели освоения дисциплины

Целями освоения дисциплины «Прикладная статистика и интеллектуальный анализ данных» являются приобретение обучающимися знаний и умений по анализу и разработке организационно-технических и экономических процессов с применением методов математического моделирования, моделированию прикладных бизнес-процессов и предметных областей в экономике.

Формируемые дисциплиной знания и умения готовят выпускника данной образовательной программы к выполнению следующей трудовой функции: С/08.6 - разработка модели бизнес-процессов заказчика (ПС 06.015 «Специалист по информационным системам»)

2. Место дисциплины в структуре ОПОП бакалавриата

Дисциплина «Прикладная статистика и интеллектуальный анализ данных» относится к обязательной части Блока 1 «Дисциплины (модули)» ОПОП.

Изучение дисциплины базируется на знаниях, умениях и готовностях, полученных в процессе изучения дисциплин образовательной программы бакалавриата по направлению подготовки 09.03.03 «Прикладная информатика»: «Математика», «Информационные технологии в профессиональной деятельности», «Основы алгоритмизации и программирования», «Теория вероятностей и математическая статистика».

В результате освоения дисциплины «Математика» студент должен знать основы математики, уметь осуществлять поиск, критический анализ информации, применять системный подход для решения поставленных задач и методы математического анализа в профессиональной деятельности.

В результате освоения дисциплины «Информационные технологии в профессиональной деятельности» студент должен знать основы информатики, современные информационные технологии и методы их использования при решении задач профессиональной деятельности, уметь использовать необходимые информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности.

В результате освоения дисциплины «Основы алгоритмизации и программирования» студент должен знать современные языки программирования и программные среды и уметь разрабатывать алгоритмы и программы, пригодные для практического применения.

В результате освоения дисциплины «Теория вероятностей и математическая статистика» студент должен знать основы теории вероятностей и математической статистики и применять методы математического моделирования в профессиональной деятельности.

Компетенции, приобретенные в ходе изучения дисциплины, могут быть использованы при освоении дисциплины «Имитационное моделирование экономических процессов», «Информационный менеджмент», «Сетевая экономика», «Теория систем и системный анализ», «Проектирование информационных систем» и последующего прохождения преддипломной практики, подготовки к итоговой государственной аттестации.

3. Результаты освоения дисциплины «Прикладная статистика и интеллектуальный анализ данных»

Процесс изучения дисциплины направлен на формирование элементов следующих компетенций в соответствии с ФГОС ВО по данному направлению:

Коды компетенции	Наименование компетенции	Индикатор достижения компетенции (закрепленный за дисциплиной)	В результате освоения дисциплины обучающийся должен:
<i>ОПК-6</i>	Способен анализировать и разрабатывать организационно-технические и экономические процессы с применением методов системного анализа и математического моделирования	<i>ОПК-6.1:</i> рассматривает основы анализа, синтеза, оценивания, математического моделирования организационно-технических и экономических процессов и систем	знать основы анализа, синтеза, оценивания, математического моделирования организационно-технических и экономических процессов и систем
		<i>ОПК-6.2:</i> применяет методы теории систем и системного анализа, математического, статистического моделирования для автоматизации организационно-технических и экономических процессов	уметь применять методы теории систем и системного анализа, математического, статистического моделирования для автоматизации организационно-технических и экономических процессов
		<i>ОПК-6.3:</i> проводит инженерные расчёты основных показателей эффективности создания и применения информационных и автоматизированных систем	владеть навыками проведения инженерных расчётов основных показателей эффективности создания и применения информационных и

			автоматизированных систем
<i>ПК-5</i>	Способен моделировать прикладные (бизнес) процессы и предметную область в экономике	<i>ПК-5.1:</i> использует методы решения прикладных задач на основе моделирования бизнес-процессов	знать методы решения прикладных задач на основе моделирования бизнес-процессов
		<i>ПК-5.2:</i> разрабатывает и документирует модели прикладных бизнес-процессов и предметной области	уметь разрабатывать и документировать модели прикладных бизнес-процессов и предметной области
		<i>ПК-5.3:</i> применяет программные средства моделирования бизнес-процессов и предметной области	владеть навыками применения программных средств моделирования бизнес-процессов и предметной области

4. Структура и содержание дисциплины «Прикладная статистика и интеллектуальный анализ данных»

4.1. Структура дисциплины

Общая трудоемкость дисциплины составляет 5 зачетных единиц, 180 часов.

№ п/п	Наименование разделов и тем дисциплины (модуля)	Семестр	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)								Формы текущего контроля успеваемости
			Контактная работа				Самостоятельная работа				
			Всего	Лекция	Лабораторные занятия	Др. виды контакт. работы	Всего	Выполнение контрольной работы	Контроль	Подготовка к экзамену	Контрольная работа
1.	Раздел 1. Прикладная статистика	6	10	4	6		77	77			
1.1.	Тема 1.1. Первичная статистическая обработка исходных данных	6	1	1			11	11			+
1.2.	Тема 1.2. Точечные и интервальные оценки параметров распределений	6					11	11			+
1.3.	Тема 1.3. Проверка статистических гипотез.	6	3	1	2		11	11			+
1.4.	Тема 1.4. Регрессионный анализ.	6					11	11			+
1.5.	Тема 1.5. Корреляционный анализ	6					11	11			+
1.6.	Тема 1.6. Дисперсионный анализ	6	3	1	2		11	11			+
1.7.	Тема 1.7. Анализ временных рядов	6	3	1	2		11	11			+
2.	Раздел 2. Интеллектуальный анализ данных	6					38,1	38,1			+
2.1.	Многомерная классификация	6					19	19			+
2.2.	Кластерный анализ данных	6					19,1	19,1			+

	<i>Подготовка к экзамену</i>									36	
	<i>Др. виды контактной работы</i>					0,9					
	<i>Контроль</i>								9		
	Общая трудоемкость, в часах		10,9	4	6	0,9	160,1	115,1	9	36	Промежуточная аттестация
											Форма Семестр
											Зачет 5
											Экзамен 5

4.2. Содержание дисциплины

4.2.1. Содержание лекционных занятий

Раздел 1. Прикладная статистика

Тема 1.1. Первичная статистическая обработка исходных данных

Методы статистики. Статистическое наблюдение и его виды. Сводка и группировка статистического материала. Вариационные ряды. Графики вариационных рядов. Эмпирические функции распределения. Основные виды распределений. Выборочные характеристики распределения. Особенности статистической обработки данных при малых выборках. Абсолютные и относительные величины. Средние величины и показатели вариации.

Тема 1.2. Точечные и интервальные оценки параметров распределений.

Оценки параметров и их свойства. Точечные и интервальные оценки. Оценивание с помощью доверительных интервалов. Доверительная вероятность и доверительный интервал. Интервальное оценивание генеральной средней величины и генеральной дисперсии.

Тема 1.3. Проверка статистических гипотез.

Основные понятия теории статистической проверки гипотез. Нулевая и альтернативная гипотезы. Критическая область. Статистика критерия. Критические значения. Ошибки первого и второго рода. Уровень значимости и мощность критерия. Проверка гипотезы о законе распределения. Критерий хи-квадрат. Критерий Колмогорова – Смирнова. Проверка гипотезы о средних величинах. Критерий Стьюдента.

Тема 1.4. Регрессионный анализ.

Постановка задачи восстановления зависимости. Оценка параметров уравнения парной регрессии. Множественная регрессия. Оценка параметров множественной регрессии. Использование уравнения регрессии для прогноза.

Тема 1.5. Корреляционный анализ.

Изучение связи между количественными переменными. Коэффициент линейной корреляции и проверка его значимости. Выборочное корреляционное отношение. Частный и множественный коэффициенты корреляции. Коэффициенты ранговой корреляции и проверка их значимости. Корреляционный анализ нечисловых данных.

Тема 1.6. Дисперсионный анализ.

Однофакторный дисперсионный анализ. Двухфакторный дисперсионный анализ с равным и неравным числом наблюдений в ячейке.

Тема 1.7. Анализ временных рядов.

Виды временных рядов (рядов динамики). Показатели абсолютного и относительного изменения рядов динамики. Индексный метод. Различные виды индексов и их свойства. Индексы Пааше и Ласпейреса. Требования, предъявляемые к индексам. Индекс Фишера. Основные методы сглаживания рядов динамики. Основные виды трендов ряда динамики. Методика выделения трендов. Выделение волн (сезонных колебаний) ряда динамики. Прогнозирование с помощью ряда динамики.

Раздел 2. Интеллектуальный анализ данных.

Тема 2.1. Многомерная классификация

Понятие о статистической классификации. Методы построения многомерных классификаций: многомерная средняя, метод шаров, метод дендритов.

Тема 2.2. Кластерный анализ

Постановка задачи кластеризации. Меры близости в кластерном анализе. Иерархические и неиерархические алгоритмы кластеризации. Представление результатов кластеризации.

4.2.2. Перечень и содержание лабораторных занятий.

№ п/п	№ раздела	Наименование лабораторных работ	Кол. часов
1	1	Проверка гипотезы о законе выборочного распределения.	2
2	1	Выявление зависимости и оценка связи между номинальными признаками.	2
3	1	Определение тренда ряда динамики.	2

5. Образовательные технологии

В процессе изучения дисциплины применяются следующие образовательные технологии:

- 1) чтение лекции с применением мультимедиа-технологий,
- 2) мастер-классы по обработке статистических данных.

В целях реализации индивидуального подхода к обучению студентов, осуществляющих учебный процесс по собственной траектории в рамках индивидуального рабочего плана, изучение данной дисциплины базируется на следующих возможностях: обеспечение внеаудиторной работы со студентами в том числе в электронной образовательной среде с использованием соответствующего программного оборудования, дистанционных форм обучения, возможностей интернет-ресурсов, индивидуальных консультаций и т.д.

6. Учебно-методическое обеспечение самостоятельной работы студентов. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины.

6.1. План самостоятельной работы студентов

Тема	Вид самостоятельной работы	Задание	Рекомендуемая литература	Кол-во часов
Тема 1.1. Первичная статистическая обработка исходных данных	Выполнение контрольной работы	Подготовить материалы по теме контрольной работы	/1/ – /5/	11
Тема 1.2. Точечные и интервальные оценки параметров распределений.	Выполнение контрольной работы	Подготовить материалы по теме контрольной работы	/1/ – /5/	11
Тема 1.3. Проверка статистических гипотез.	Выполнение контрольной работы	Подготовить материалы по теме контрольной работы	/1/ – /5/	11

Тема 1.4. Регрессионный анализ.	Выполнение контрольной работы	Подготовить материалы по теме контрольной работы	/1/ – /5/	11
Тема 1.5. Корреляционный анализ.	Выполнение контрольной работы	Подготовить материалы по теме контрольной работы	/1/ – /5/	11
Тема 1.6. Дисперсионный анализ.	Выполнение контрольной работы	Подготовить материалы по теме контрольной работы	/1/ – /5/	11
Тема 1.7. Анализ временных рядов.	Выполнение контрольной работы	Подготовить материалы по теме контрольной работы	/1/ – /5/	11
Тема 2.1. Многомерная классификация.	Выполнение контрольной работы	Подготовить материалы по теме контрольной работы	/1/ – /5/	19
Тема 2.2. Кластерный анализ	Выполнение контрольной работы	Подготовить материалы по теме контрольной работы	/1/ – /5/	19,1
Темы 1.1-1.7 Темы 2.1-2.2	Подготовка к экзамену	Подготовить ответы на теоретические вопросы и практические задания	/1/ – /5/	36

6.2. Методические указания по организации самостоятельной работы студентов.

Самостоятельная работа по дисциплине «Прикладная статистика интеллектуальный анализ данных» включает:

- работу с учебной литературой и ресурсами сети Интернет при выполнении контрольной работы
- подготовку к экзамену

В ходе самостоятельной работы студенты должны ориентироваться на список источников, которые рекомендованы преподавателем, а также самостоятельно определять источники получения информации (печатные и электронные издания, электронные ресурсы).

6.3. Материалы для проведения текущего и промежуточного контроля знаний студентов

Контроль освоения компетенций

№ п\п	Вид контроля	Контролируемые темы (разделы)	Компетенции, компоненты которых контролируются
1	Текущий: контрольная работа	Разделы 1, 2	ОПК-6, ПК-5
2	Промежуточный: зачёт	Разделы 1, 2	ОПК-6, ПК-5
3	Промежуточный: экзамен	Разделы 1, 2	ОПК-6, ПК-5

Материалы для проведения текущего контроля знаний и промежуточной аттестации составляют отдельный документ – Фонд оценочных средств по дисциплине «Прикладная статистика и интеллектуальный анализ данных».

Демонстрационные варианты оценочных средств для каждого вида контроля смотри <http://moodle.pnzgu.ru/enrol/index.php?id=48794>.

7. Учебно-методическое и материально-техническое обеспечение дисциплины «Прикладная статистика и интеллектуальный анализ данных»

а) учебная литература:

1. Общая и прикладная статистика: Учеб. для студ. высш. проф. обр./Р.Н.Пахунова, П.Ф.Аскеров и др.; Под общ. ред. Р.Н.Пахуновой - М.: НИЦ ИНФРА-М, 2013-272с. + (Доп. мат. znanium.com) - (ВО:Бакалавр.) [Электронный ресурс]. Режим доступа:<http://znanium.com/bookread2.php?book=404310>

2. Статистический анализ данных в MS Excel: Учебное пособие/Козлов А.Ю., Мхитарян В.С., Шишов В.Ф. - М.:НИЦ ИНФРА-М, 2016. - 320 с. – (Высшее образование:Бакалавриат) [Электронный ресурс]. Режим доступа:<http://znanium.com/bookread2.php?book=558444>

б) Интернет-ресурсы

3. НОУ Интуит. Учебный курс. Прикладная статистика. [Электронный ресурс]. Режим доступа:<http://www.intuit.ru/studies/courses/546/402/info>

4. НОУ Интуит. Учебный курс. DataMining.[Электронный ресурс]. Режим доступа: <http://www.intuit.ru/studies/courses/6/6/info>

б) Интернет-ресурсы

5. Портал Федеральной службы государственной статистики. [Электронный ресурс].режим доступа:<http://www.gks.ru>

в) программное обеспечение

Microsoft Office Excel, R, SPSS, Statistica.

г) другое материально-техническое обеспечение

Для проведения лекционных занятий необходима аудитория, оснащённая ноутбуком, компьютерным проектором, проекционным экраном, шторами, сетью электропитания 220В.

Для проведения лабораторных занятий используется компьютерный класс, оборудованный локальной сетью и выходом в Internet.

Рабочая программа дисциплины «Прикладная статистика и интеллектуальный анализ данных» составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 09.03.03 «Прикладная информатика», утвержденного приказом Министерства образования и науки Российской Федерации от «19» сентября 2017 г. № 922.

Программу составил:

Буданов К.М., старший преподаватель каф. ИВС

(Ф.И.О., должность, подпись)

Настоящая программа не может быть воспроизведена ни в какой форме без предварительного письменного разрешения кафедры-разработчика программы.

Программа одобрена на заседании кафедры _____

Протокол № ____ от « ____ » _____ 20__ года

Зав. кафедрой _____

(подпись, Ф.И.О.)

Программа одобрена методической комиссией _____ факультета
(института)

Протокол № ____ от « ____ » _____ 20__ года

Председатель методической комиссии
_____ факультета (института)

(подпись)

(Ф.И.О.)

Сведения о переутверждении программы на очередной учебный год и регистрации изменений

Учебный год	Решение кафедры (№ протокола, дата)	Внесенные изменения	Подпись зав. кафедрой