

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ПЕНЗЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ

УТВЕРЖДАЮ
Декан ФВТ _____ Л.Р. Фионова
« 16 » _____ 2015 г.



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Б1.2.2 Прикладная статистика и интеллектуальный анализ данных

Направление подготовки – *09.03.03 Прикладная информатика*

Профиль подготовки – *Прикладная информатика в экономике*

Квалификация (степень) выпускника – *бакалавр*

Форма обучения – *заочная*

г. Пенза, 2015 г.

1. Цели освоения дисциплины

Целями освоения дисциплины «Прикладная статистика и интеллектуальный анализ данных» являются ознакомление студентов с основными статистическими методами, применяемыми при анализе данных в экономике, а также подготовка к самостоятельному решению задач статистического анализа данных, начиная от их формулирования исходных задач соответствующей предметной области на языке прикладной статистики, выбора методов решения и критериев качества полученных решений и заканчивая формулировкой полученных выводов на языке предметной области и интерпретацией полученных результатов анализа.

2. Место дисциплины в структуре ОПОП бакалавриата

Учебная дисциплина «Прикладная статистика и интеллектуальный анализ данных» относится к вариативной части ОПОП и изучается в 2 семестре.

Для изучения дисциплины «Прикладная статистика и интеллектуальный анализ данных» необходимы:

– знание и умения по применению различных математических методов решения задач, полученные в ходе изучения курса «Математика» (Б1.1.5) в 1 и 2 семестрах;

– знания и умения по применению информационных технологий обработки числовой информации, полученные в ходе изучения курса «Информационные технологии в профессиональной деятельности» (Б1.1.4) в 1 семестре;

– знания и умения по разработке алгоритмов решения задач, полученные в ходе изучения курса «Основы алгоритмизации и программирования» (Б1.1.11) в 2 семестре;

Теоретические знания и практические навыки, полученные при изучении дисциплины «Прикладная статистика и интеллектуальный анализ данных», предшествуют изучению следующих дисциплин: «Основы банковской деятельности» (Б1.2.11), «Математическая экономика» (Б1.2.5), «Эконометрика» (Б1.2.10). Эти знания и практические навыки также необходимы для выполнения курсовых работ и выпускной бакалаврской работы.

Минимальные требования к входным знаниям умениям, необходимым для успешного усвоения данной дисциплины – удовлетворительное усвоение программ по разделам дисциплин «Математика», «Информационные технологии в профессиональной деятельности», «Основы алгоритмизации и программирования».

3. Компетенции обучающегося, формируемые в результате 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	Структурные элементы компетенции (в результате освоения дисциплины обучающийся должен знать, уметь, владеть)
ОПК-2	способен анализировать социально-экономические задачи и процессы с применением методов системного анализа и математического моделирования	Знать: методы анализа статистических зависимостей, методы построения и анализа моделей регрессии, методы анализа временных рядов и дисперсионного анализа, методы классификации и типизации неоднородных объектов
		Уметь: выбирать методы обработки данных и критерии качества полученных решений, формулировать полученные выводы на языке предметной области
		Владеть: современными программными средствами для решения задач анализа статистических данных
ПСК-1	способен проводить анализ, прогнозирование, моделирование и реинжиниринг экономических процессов с применением современных ИКТ	Знать: основные методы анализа и прогнозирования, применяемые для исследования экономических процессов
		Уметь: применять методы анализа и прогнозирования для решения поставленных задач и интерпретировать полученные решения
		Владеть: современными программными средствами для решения задач анализа и прогнозирования экономических процессов
ПК-6	способен собирать детальную информацию для формализации требований пользователей заказчика	Знать: методику сбора статистических данных
		Уметь: проводить обработку статистических данных
		Владеть: современными программными средствами и информационными технологиями обработки статистических данных

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

4.1. Структура дисциплины

Общая трудоемкость дисциплины составляет 4 зачетных единицы, 144 часа.

№ п/п	Наименование разделов и тем дисциплины (модуля)	Семестр	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)						Формы текущего контроля успеваемости
			Аудиторная работа			Самостоятельная работа			
			Всего	Лекция	Лабораторные занятия	Всего	Выполнение контрольной работы	Подготовка к зачёту	
1	Раздел 1. Прикладная статистика	2	14	4	10	128	96	32	Защита контрольной работы
1.1.	Тема 1.1. Первичная статистическая обработка исходных данных	2	3	1	2	16	12	4	Защита контрольной работы
1.2	Тема 1.2. Точечные и интервальные оценки параметров распределений.	2	2		2	16	12	4	Защита контрольной работы
1.3	Тема 1.3. Проверка статистических гипотез.	2	3	1	2	16	12	4	Защита контрольной работы
1.4	Тема 1.4. Регрессионный анализ.	2				16	12	4	Защита контрольной работы
1.5	Тема 1.5. Корреляционный анализ.	2	3	1	2	16	12	4	Защита контрольной работы
1.6	Тема 1.6. Дисперсионный анализ.	2				16	12	4	Защита контрольной работы

1.7	Тема 1.7. Анализ временных рядов.	2	3	1	2	16	12	4	Защита контрольной работы
1.8	Тема 1.8. Многомерная классификация.	2				16	12	4	Защита контрольной работы
2	Раздел 2. Интеллектуальный анализ данных	2				16	12	4	Защита контрольной работы
2.1	Тема 2.1 OLAP-системы.	2				16	12	4	Защита контрольной работы
	Общая трудоемкость, в часах		14	4	10	144	108	36	Промежуточная аттестация
									Форма
									Семестр
									зачёт
									2

4.2. Содержание дисциплины

4.2.1. Содержание лекционного курса

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. Оценки 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.

Тема 1.8. Многомерная классификация

Понятие о статистической классификации. Методы построения многомерных классификаций: многомерные средние, кластерный анализ.

Раздел 2. Интеллектуальный анализ данных.

Тема 2.1. OLAP-системы

Многомерная модель данных. Определение OLAP-систем. Концептуальное многомерное представление данных. Архитектура OLAP-систем: MOLAP, ROLAP, HOLAP.

4.2.2. Перечень и содержание лабораторных занятий.

№ п/п	№ раздела	Наименование лабораторных работ	Кол. часов
1	1	Формирование выборки с заданным законом распределения. Построение вариационного ряда.	2
2	1	Оценка параметров генеральной совокупности по выборочному распределению	2
3	1	Проверка гипотезы о законе выборочного распределения	2
4	1	Выявление зависимости и оценка связи между номинальными признаками	4
5	1	Определение тренда ряда динамики	4

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 Microsoft Excel

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	а) подготовка к зачёту	ответить на вопросы из списка вопросов к зачёту	/1/ – /5/	4
	б) выполнение контрольной работы	подготовить материалы по теме контрольной работы	/1/ – /5/	12
Тема 1.2. Точечные и интервальные оценки параметров распределений.	а) подготовка к зачёту	ответить на вопросы из списка вопросов к зачёту	/1/ – /5/	4
	б) выполнение контрольной работы	подготовить материалы по теме контрольной работы	/1/ – /5/	12
Тема 1.3. Проверка статистических гипотез.	а) подготовка к зачёту	ответить на вопросы из списка вопросов к зачёту	/1/ – /5/	4
	б) выполнение контрольной работы	подготовить материалы по теме контрольной работы	/1/ – /5/	12
Тема 1.4. Регрессионный анализ.	а) подготовка к зачёту	ответить на вопросы из списка вопросов к зачёту	/1/ – /5/	4
	б) выполнение контрольной работы	подготовить материалы по теме контрольной работы	/1/ – /5/	12
Тема 1.5. Корреляционный анализ.	а) подготовка к зачёту	ответить на вопросы из списка вопросов к зачёту	/1/ – /5/	4
	б) выполнение контрольной работы	подготовить материалы по теме контрольной работы	/1/ – /5/	12

Тема 1.6. Дисперсионный анализ.	а) подготовка к зачёту	ответить на вопросы из списка вопросов к зачёту	/1/ – /5/	4
	б) выполнение контрольной работы	подготовить материалы по теме контрольной работы	/1/ – /5/	12
Тема 1.7. Анализ временных рядов.	а) подготовка к зачёту	ответить на вопросы из списка вопросов к зачёту	/1/ – /5/	4
	б) выполнение контрольной работы	подготовить материалы по теме контрольной работы	/1/ – /5/	12
Тема 1.8. Многомерная классификация.	а) подготовка к зачёту	ответить на вопросы из списка вопросов к зачёту	/1/ – /5/	4
	б) выполнение контрольной работы	подготовить материалы по теме контрольной работы	/1/ – /5/	12
Тема 2.1. OLAP-системы	а) подготовка к зачёту	ответить на вопросы из списка вопросов к зачёту	/1/ – /5/	4
	б) выполнение контрольной работы	подготовить материалы по теме контрольной работы		12

6.2. Методические указания по организации самостоятельной работы студентов.

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

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

При проведении 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	Контрольная работа	Разделы 1 – 2	ОПК-2, ПСК-1, ПК-6
2	Зачёт	Разделы 1 – 2	ОПК-2, ПСК-1, ПК-6

Тематика контрольных работ

1. Виды и методы статистического наблюдения.
2. Методика оценивания параметров генеральной совокупности по выборочным данным.
3. Проверка гипотез в статистике.
4. Постановка задачи восстановления зависимости и методика её решения.
5. Оценка параметров уравнения парной регрессии.
6. Коэффициент линейной корреляции: назначение, методика расчёта, проверка значимости.
7. Выборочное корреляционное отношение: назначение, методика расчёта, основные свойства.
8. Частные и множественный коэффициенты корреляции и методика их расчёта.
9. Методика одномерного и многомерного дисперсионного анализа.
10. Методика дисперсионного анализа с повторными измерениями.
11. Методика дисперсионного анализа моделей с постоянными факторами.
12. Методика дисперсионного анализа моделей со случайными факторами.
13. Методика дисперсионного анализа смешанных моделей.
14. Виды индексов в статистике и их практическое применение.
15. Выделение сезонных колебаний с помощью спектрального анализа временных рядов.
16. Авторегрессионные модели.
17. Методы анализа и прогнозирования временных рядов.
18. Ряды Фурье и их применение для выравнивания рядов динамики.
19. Методика компонентного анализа.
20. Методика факторного анализа.
21. Методика дискриминантного анализа.
22. Методика кластерного анализа.
23. Методика многомерного шкалирования.
24. Технология Data Mining в интеллектуальном анализе данных.

Вопросы к зачёту

1. Эмпирические функции распределения и основные виды статистических распределений.
2. Основные выборочные характеристики распределения.
3. Особенности статистической обработки данных при малых выборках.
4. Виды абсолютных величин в статистике.
5. Виды относительных величин в статистике.
6. Виды средних величин в статистике.
7. Абсолютные и относительные показатели вариации.
8. Методика интервального оценивания генеральной средней величины и генеральной дисперсии.
9. Проверка гипотезы о законе распределения с помощью критерия хи-квадрат.
10. Критерий Колмогорова – Смирнова и его практического применение.
11. Проверка гипотезы о средних величинах.
12. Критерий Стьюдента и область его применения.
13. Понятие о множественной регрессии и методика оценки параметров множественной регрессии.
14. Формирование прогноза на основе уравнения регрессии.
15. Коэффициенты ранговой корреляции и проверка их значимости.
16. Методика корреляционного анализа нечисловых данных.
17. Методика однофакторного дисперсионного анализа.
18. Методика двухфакторного дисперсионного анализа.
19. Ряды динамики и показатели их абсолютного и относительного изменения.
20. Основные методы сглаживания ряда динамики.
21. Тренды ряда динамики и методика их выделения.
22. Методика выделения волн (сезонных колебаний) ряда динамики.
23. Построение прогноза по ряду динамики.
24. Методы построения многомерных классификаций: многомерные средние, кластерный анализ.
25. Концепция многомерного представления данных.
26. OLAP-системы и их архитектура: MOLAP, ROLAP, HOLAP.

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

а) основная 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>

б) дополнительная литература

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>

в) 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. Data Mining. [Электронный ресурс]. Режим доступа: <http://www.intuit.ru/studies/courses/6/6/info>

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

г) программное обеспечение

Microsoft Office Excel

8. Материально-техническое обеспечение дисциплины

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

Для проведения лабораторных занятий используется дисплейный класс, оснащённый персональными компьютерами, на которых установлено программное средство Microsoft Office Excel.

Рабочая программа дисциплины «Прикладная статистика и интеллектуальный анализ данных» составлена в соответствии с требованиями ФГОС ВО по направлению 09.03.03 «Прикладная информатика».

Программу составил:

1. ст.преподаватель каф. ИВС


(подпись)

К.М. Буданов

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

Программа одобрена на заседании кафедры «Информационно-вычислительные системы»

Протокол № 12 от 15.06.2015 года

Зав. кафедрой ИВС


(подпись)

Ю.Н. Косников

Программа одобрена методической комиссией ФВТ

Протокол № 6 от «15» 06 2016 года

Председатель методической комиссии ФВТ


(подпись)

Н.Н. Коннов

Сведения о переутверждении программы на очередной учебный год и регистрации изменений

Учебный год	Решение кафедры (№ протокола, дата, подпись зав. кафедрой)	Внесенные изменения	Номера листов (страниц)		
			замененных	новых	аннулированных
2016/2017	Прот. № 11 от 22.06.16г. <i>[подпись]</i>	Переутверждено без изменений			
2017/2018	Прот. № 14 от 27.06.2017г. <i>[подпись]</i>	Введена информация из ФЭС	14		
2018/2019	Прот. № 14 от 27.06.2018 <i>[подпись]</i>	Переутверждено без изменений			