

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ОБРАЗОВАНИЯ  
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИНЖЕНЕРНЫХ ТЕХНОЛОГИЙ»**

**УТВЕРЖДАЮ**

Проректор по учебной работе

\_\_\_\_\_ Василенко В.Н.  
(подпись) (ф.и.о.)

«25» мая 2023 г.

**РАБОЧАЯ ПРОГРАММА  
ДИСЦИПЛИНЫ**

**Прикладная статистика**

Направление подготовки

09.03.03 Прикладная информатика

Направленность (профиль) подготовки

**Цифровизация бизнес-процессов**

Квалификация выпускника

**Бакалавр**

Воронеж

## 1. Цели и задачи дисциплины

Целью освоения дисциплины «Прикладная статистика» является освоение компетенций обучающегося в области профессиональной деятельности по Реестру Минтруда – Об Связь, информационные и коммуникационные технологии и сфере профессиональной деятельности – Сфера исследования, разработки, внедрения и сопровождения информационных технологий и систем:

В рамках освоения ОП ВО выпускники готовятся к решению задач профессиональной деятельности следующих типов:

- производственно-технологический;
- организационно-управленческий;
- проектный.

Программа составлена в соответствии с требованиями Федерального государственного образовательного стандарта высшего образования, на основе примерной основной образовательной программы высшего образования по направлению подготовки 09.03.03 «Прикладная информатика», (уровень образования - бакалавриат).

## 2. Перечень планируемых результатов обучения, соотнесенных с планируемыми результатами освоения образовательной программы

№ п/п	Код компетенции	Формулировка компетенции	Код и наименование индикатора достижения компетенции
1	УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	ИД1 <sub>УК-1</sub> Умеет анализировать и систематизировать разнородные данные, оценивать эффективность процедур анализа проблем и принятия решений в профессиональной деятельности.
2	УК-2	Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений	ИД1 <sub>УК-2</sub> Умеет анализировать альтернативные варианты решений для достижения намеченных результатов; разрабатывать план, определять целевые этапы и основные направления работ.
			ИД2 <sub>УК-2</sub> Владеет методиками разработки цели и задач проекта; методами оценки продолжительности и стоимости проекта, а также потребности в ресурсах.
3	ПКВ-4	Способность составлять технико-экономическое обоснование проектных решений и техническое задание на разработку информационной системы	ИД1 <sub>ПКВ-4</sub> Умеет составлять технико-экономического обоснования проектных решений и техническое задание на разработку ИС
4	ПКВ-7	Способность настраивать, эксплуатировать и сопровождать информационные системы и сервисы	ИД1 <sub>ПКВ-7</sub> Участвует в экспертном тестировании ИС на этапе опытной эксплуатации

Код и наименование индикатора достижения компетенции	Результаты обучения (показатели оценивания)
ИД1 <sub>УК-1</sub> Умеет анализировать и систематизировать разнородные данные, оценивать эффективность процедур анализа проблем и принятия решений в профессиональной деятельности.	Знает: методы и принципы сбора разнородных данных
	Умеет: выполнять систематизацию разнородных данных
	Имеет навыки: анализа данных и оценки эффективности процедур принятия решений

ИД1 <sub>ук-2</sub> Умеет анализировать альтернативные варианты решений для достижения намеченных результатов; разрабатывать план, определять целевые этапы и основные направления работ.	Знает: методики составления альтернативных вариантов
	Умеет: составлять альтернативные варианты решения задач
	Имеет навыки: анализа вариантов решения задачи
ИД2 <sub>ук-2</sub> Владеет методиками разработки цели и задач проекта; методами оценки продолжительности и стоимости проекта, а также потребности в ресурсах.	Знает: методики оценки потребности в ресурсах
	Умеет: применять соответствующие методы оценки потребности в ресурсах
	Имеет навыки: оценивания проектов с точки зрения потребности в ресурсах
ИД1 <sub>пкв-4</sub> Умеет составлять технико-экономического обоснования проектных решений и техническое задание на разработку ИС	Знает: приемы и методы составления технико-экономического обоснования
	Умеет: составлять планы проведения технико-экономического обоснования
	Имеет навыки: выполнения технико-экономического обоснования проектных решений
ИД1 <sub>пкв-7</sub> Участвует в экспертном тестировании ИС на этапе опытной эксплуатации	Знает: приемы и метода проведения тестирования систем
	Умеет: применять соответствующие приемы и методы тестирования систем
	Имеет навыки: проведения тестирования информационных систем

### 3. Место дисциплины в структуре ОП ВО

Дисциплина Прикладная статистика относится к блоку Б1 и ее части: дисциплины по выбору Блока 1ООП.

Изучение дисциплины «Прикладная статистика» опирается на знания, полученные при обучении в школе, а также по дисциплинам «Математика», «Вероятностное моделирование информационных и экономических процессов», «Теория вероятностей и математическая статистика».

Дисциплина «Прикладная статистика» является предшествующей для освоения дисциплин профессиональной направленности.

### 4. Объем дисциплины и виды учебных занятий

Общая трудоемкость дисциплины составляет 3 зачетных единицы.

Виды учебной работы	Всего акад. часов	Распределение трудоемкости по семестрам
Общая трудоемкость дисциплины	<b>108</b>	<b>108</b>
<b>Контактная работа,</b> <i>в т.ч. аудиторные занятия:</i>	<b>45,85</b>	<b>45,85</b>
Лекции	15	15
Практические занятия (ПЗ)	15	15
<i>в т.ч. в форме практической подготовки</i>	30	30
Консультации текущие	30	30
Виды аттестации (зачет, экзамен)	0,75	0,75
<b>Самостоятельная работа:</b>	0,1	0,1
Проработка материалов по конспекту лекций (собеседование, тестирование)	<b>62,15</b>	<b>62,15</b>
Проработка материалов по учебнику (собеседование, тестирование)	7,65	7,65
Подготовка к коллоквиуму	22,5	22,5
Выполнение расчетов для практических работ	5	5
Подготовка к аудиторной контрольной работе	5	5
Выполнение расчетов для РПР	7	7
Оформление текста пояснительной записки для РПР	10	10

**5 Содержание дисциплины, структурированное по разделам с указанием отведенного на них количества академических часов и видов учебных занятий**

**5.1 Содержание разделов дисциплины**

№ п/п	Наименование раздела дисциплины	Содержание раздела (указываются темы и дидактические единицы)	Трудоемкость раздела, ак. часы
5 семестр			
1	Выбор способов решения задач в рамках поставленной цели. Язык R	Обзор пакетов статистической обработки данных с точки зрения опытной эксплуатации информационных систем и потребности в ресурсах.	22
		Установка и экосистема R. Сбор и систематизация разнородных данных в R. Анализ данных посредством построения графиков в R.	
2	Проверка статистических гипотез как оптимальный способ решения для достижения намеченных результатов	Случайные величины и работа с ними в R. Подбор функции распределения по эмпирическим данным. Виды функции Пирсона. Параметрические и непараметрические критерии.	29
		Способы решения задачи о проверке гипотез о распределении. Описание базы данных. Описательные статистики. Тесты на симметричность, нормальность распределения, соответствие распределений, проверку гипотез	
3	Регрессионный анализ данных с целью оценки эффективности принятия решений в профессиональной деятельности	Линейная регрессия. Постановка задачи. Метод наименьших квадратов. Нахождение уравнения регрессии. Оценка тесноты связи. Условия Гаусса-Маркова. Доверительные интервалы. Проверка значимости. Преобразования к линейному виду	17
		Множественная линейная регрессия. Расчет параметров. Проверка достоверности модели как прием тестирования системы.	
4	Методы отбора признаков как один из вариантов решения задачи для достижения намеченных результатов	Регрессия от всех переменных. Регрессия по всем подмножествам. Step-wise регрессия. Случайный лес для отбора признаков. Регрессия с регуляризацией. Метод вычисления процентной важности переменных. MARS-метод.	17,65
5	Использование методов предиктивной аналитики при составлении технико-экономического обоснования проектных решений	Анализ временных рядов как метод экспертного тестирования информационной системы на этапе опытной эксплуатации	21,5
		<i>Консультации текущие</i>	<i>0,75</i>
		<i>Зачет</i>	<i>0,1</i>

## 5.2 Разделы дисциплины и виды занятий

№ п/п	Наименование раздела дисциплины	Лекции, ак. ч	ПЗ, ак. ч	СРО, ак. ч
5 семестр				
1	Выбор способов решения задач в рамках поставленной цели. Язык R	3	8	11
2	Проверка статистических гипотез как оптимальный способ решения для достижения намеченных результатов	3	6	20
3	Регрессионный анализ данных с целью оценки эффективности принятия решений в профессиональной деятельности	3	6	8
4	Методы отбора признаков как один из вариантов решения задачи для достижения намеченных результатов	3	6	8,65
5	Использование методов предиктивной аналитики при составлении технико-экономического обоснования проектных решений	3	4	14,5
	<i>Консультации текущие</i>	<i>0,75</i>		
	<i>Зачет</i>	<i>0,1</i>		

### 5.2.1 Лекции

№ п/п	Наименование раздела дисциплины	Тематика лекционных занятий	Трудоемкость, ак. ч
1	Выбор способов решения задач в рамках поставленной цели. Язык R	Обзор пакетов статистической обработки данных с точки зрения опытной эксплуатации информационных систем и потребности в ресурсах.	3
		Установка и экосистема R. Сбор и систематизация разнородных данных в R. Анализ данных посредством построения графиков в R.	
2	Проверка статистических гипотез как оптимальный способ решения для достижения намеченных результатов	Случайные величины и работа с ними в R. Подбор функции распределения по эмпирическим данным. Виды функции Пирсона. Параметрические и непараметрические критерии.	3
		Способы решения задачи о проверке гипотез о распределении. Описание базы данных. Описательные статистики. Тесты на симметричность, нормальность распределения, соответствие распределений, проверку гипотез	
3	Регрессионный анализ данных с целью оценки эффективности принятия решений в профессиональной деятельности	Линейная регрессия. Постановка задачи. Метод наименьших квадратов. Нахождение уравнения регрессии. Оценка тесноты связи. Условия Гаусса-Маркова. Доверительные интервалы. Проверка значимости. Преобразования к линейному виду	3
		Множественная линейная регрессия. Расчет параметров. Проверка достоверности модели как прием тестирования системы.	
4	Методы отбора признаков как один из вариантов решения задачи для достижения намеченных результатов	Регрессия от всех переменных. Регрессия по всем подмножествам. Step-wise регрессия. Случайный лес для отбора признаков. Регрессия с регуляризацией. Метод вычисления процентной важности переменных. MARS-метод.	3
5	Использование методов предиктивной аналитики при составлении технико-экономического обоснования	Анализ временных рядов как метод экспертного тестирования информационной системы на этапе опытной эксплуатации	3

ния проектных решений		
-----------------------	--	--

### 5.2.2 Практические занятия

№ п/п	Наименование раздела дисциплины	Тематика практических занятий	Трудоемкость, ак. час
1	Выбор способов решения задач в рамках поставленной цели. Язык R	Установка R*	2
2		Основы работы с R*	2
3		Работа с базами данных в R*	2
4		Графики в R*	2
5	Проверка статистических гипотез как оптимальный способ решения для достижения намеченных результатов	Подбор функции распределения по эмпирическим данным*	2
6		Параметрические и непараметрические критерии*	2
7		Проведение тестов на симметричность, нормальность распределения, соответствие распределений, проверку гипотез*	2
8	Регрессионный анализ данных с целью оценки эффективности принятия решений в профессиональной деятельности	Построение уравнений парной линейной регрессии*	2
9		Построение уравнений множественной линейной регрессии*	2
10		Проверка достоверностей моделей*	2
11	Методы отбора признаков как один из вариантов решения задачи для достижения намеченных результатов	Построение поля корреляции*	2
12		Создание переменных-функционалов от исходных данных*	2
13		Исследование моделей с различными признаками*	2
14	Использование методов предиктивной аналитики при составлении технико-экономического обоснования проектных решений	Разделение рядов на обучающую и тестовую выборки*	2
15		Построение моделей и расчет метрик с целью выбора оптимальной с точки зрения качества модели *	2

\*в форме практической подготовки

### 5.2.3 Лабораторный практикум - не предусмотрен

### 5.2.4 Самостоятельная работа обучающихся (СРО)

№ п/п	Наименование раздела дисциплины	Вид СРО	Трудоемкость, час
7 семестр			
1	Выбор способов решения задач в рамках поставленной цели. Язык R	Проработка конспекта лекций	2
		Проработка материалов по учебнику	5
		Подготовка к коллоквиуму	2
		Выполнение расчетов для практических работ	2
2	Проверка статистических гипотез как оптимальный способ решения для достижения намеченных результатов	Проработка конспекта лекций	2
		Проработка материалов по учебнику	5
		Подготовка к коллоквиуму	3
		Выполнение расчетов и оформление текста пояснительной записки для РПР	8
		Выполнение расчетов для практических работ	2
3	Регрессионный анализ данных с целью оценки эффективности принятия решений в профессиональной деятельности	Проработка конспекта лекций	2
		Проработка материалов по учебнику	5
		Выполнение расчетов для практических работ	1
4	Методы отбора признаков как один из вариантов решения задачи для достижения намеченных результатов	Проработка конспекта лекций	1,65
		Подготовка к аудиторной контрольной работе	7
5	Использование методов предик-	Проработка материалов по учебнику	7,5

тивной аналитики при составлении технико-экономического обоснования проектных решений	Выполнение расчетов для РПР	5
	Оформление текста пояснительной записки для РПР	2

## 6 Учебно-методическое и информационное обеспечение дисциплины

### 6.1 Основная литература

1. Прикладная математическая статистика: учебное пособие - Томск: ТУСУР, 2016 – Режим доступа: <http://biblioclub.ru/index.php?page=book&id=480889>
2. Горяинова Е. Р., Панков А. Р., Платонов Е. Н. Прикладные методы анализа статистических данных: учебное пособие - Москва: Издательский дом Высшей школы экономики, 2012 – Режим доступа: <http://biblioclub.ru/index.php?page=book&id=227280>
3. Прикладная математическая статистика / сост. А.А. Мицель ; Министерство образования и науки Российской Федерации. – Томск : ТУСУР, 2016. – 113 с. : ил. – Режим доступа: по подписке. – RL: <http://biblioclub.ru/index.php?page=book&id=480889>

### 6.2 Дополнительная литература

1. Мещеряков В. В. Задачи по статистике и регрессионному анализу с MATLAB - Москва: Диалог-МИФИ, 2009 – Режим доступа: <http://biblioclub.ru/index.php?page=book&id=136083>
2. Кобзарь А. И. Прикладная математическая статистика - Москва: Физматлит, 2012 – Режим доступа: <http://biblioclub.ru/index.php?page=book&id=82617>
3. Орлов А. И. Прикладная статистика - Москва: Интернет-Университет Информационных Технологий, 2009 – Режим доступа: <http://biblioclub.ru/index.php?page=book&id=234537>
4. Кельберт М. Я., Сухов Ю. М. Вероятность и статистика в примерах и задачах. Т. 2. Марковские цепи как отправная точка теории случайных процессов - Москва: МЦНМО, 2010 – Режим доступа: <http://biblioclub.ru/index.php?page=book&id=63156>
5. Зехин В. А., Мхитарян В. С., Айвазян С. А. Практикум по многомерным статистическим методам: учебное пособие - Москва: Московский государственный университет экономики, статистики и информатики, 2003 – Режим доступа: <http://biblioclub.ru/index.php?page=book&id=90409>
6. Путко Б. А., Кремер Н. Ш. Эконометрика: учебник - Москва: Юнити-Дана, 2012 – Режим доступа: <http://biblioclub.ru/index.php?page=book&id=118251>
7. Балдин К. В., Быстров О. Ф., Соколов М. М. Эконометрика: учебное пособие - Москва: Юнити-Дана, 2015– Режим доступа: <http://biblioclub.ru/index.php?page=book&id=114533>
8. Садовникова Н. А., Шмойлова Р. А. Анализ временных рядов и прогнозирование: учебно-методический комплекс. Вып. 5 - Москва: Евразийский открытый институт, 2011– Режим доступа: <http://biblioclub.ru/index.php?page=book&id=90649>

### 6.3 Перечень учебно-методического обеспечения для самостоятельной работы обучающихся

1. Богачев, Б. М. Задачи по теории вероятностей [Электронный ресурс] : учебное пособие к практическим занятиям и выполнению контрольных работ / Б. М. Богачев, В. И. Ряжских, В. В. Сысоев; ВГУИТ, Кафедра информационных технологий, моделирования и управления. - Воронеж, 2012. - 124 с.

### 6.4 Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. Сайт научной библиотеки ВГУИТ <<http://cnit.vsuet.ru>>.

2. Базовые федеральные образовательные порталы. <[http://www.edu.ru/db/portal/sites/portal\\_page.htm](http://www.edu.ru/db/portal/sites/portal_page.htm)>.

3. Государственная публичная научно-техническая библиотека. <[www.gpntb.ru/](http://www.gpntb.ru/)>.

4. Информационно-коммуникационные технологии в образовании. Система федеральных образовательных порталов. <<http://www.ict.edu.ru/>>.

5. Национальная электронная библиотека. <[www.nns.ru/](http://www.nns.ru/)>..

6. Поисковая система «Рамблер». <[www.rambler.ru/](http://www.rambler.ru/)>.

7. Поисковая система «Yahoo». <[www.yahoo.com/](http://www.yahoo.com/)>.

8. Поисковая система «Яндекс». <[www.yandex.ru/](http://www.yandex.ru/)>.

9. Российская государственная библиотека. <[www.rsl.ru/](http://www.rsl.ru/)>.

10. Российская национальная библиотека. <[www.nlr.ru/](http://www.nlr.ru/)>.

## **6.5 Методические указания для обучающихся по освоению дисциплины**

Методические указания для обучающихся по освоению дисциплин (модулей) в ФГБОУ ВО ВГУИТ [Электронный ресурс] : методические указания для обучающихся на всех уровнях высшего образования / М. М. Данылиев, Р. Н. Плотникова; ВГУИТ, Учебно-методическое управление. - Воронеж : ВГУИТ, 2015. – Режим доступа: <http://biblos.vsu.ru/ProtectedView/Book/ViewBook/2488>. - Загл. с экрана.

## **6.6 Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем**

*Программное обеспечение: ОС Windows; MSOffice; RStudio*

*Локальная сеть университета и глобальная сеть Internet.*

## **7 Материально-техническое обеспечение дисциплины**

Для практических работ: ауд. 336а - компьютерный класс каф. ВМиИТ

Количество ПЭВМ – 9 (Core i3 540) Microsoft Windows Professional 7 Russian Upgrade Academic OPEN 1 License No Level #47881748 от 24.12.2010 г. <http://eopen.microsoft.com>;

Microsoft Office 2007, Microsoft Office Professional Plus 2007 (Visio, Project) Russian Academic OPEN No Level #44822753 от 17.11.2008 <http://eopen.microsoft.com>;

Для лекционных занятий используется лекционный аудиторный фонд университета и переносное мультимедийное оборудование – ноутбук и экран.

## **8 Оценочные материалы для промежуточной аттестации обучающихся по дисциплине**

8.1 Оценочные материалы (ОМ) для дисциплины включают в себя:

- перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы;

- описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания;

- типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы;

- методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций.

8.2 Для каждого результата обучения по дисциплине (модулю) определяются показатели и критерии оценивания сформированности компетенций на различных этапах их формирования, шкалы и процедуры оценивания.



ОМ представляются в виде отдельного документа и входят в состав рабочей программы дисциплины (модуля) в виде приложения.

Оценочные материалы формируются в соответствии с П ВГУИТ «Положение об оценочных материалах»

**ПРИЛОЖЕНИЕ**  
**к рабочей программе**

**1. Организационно-методические данные дисциплины для очно-заочной или заочной форм обучения**

**1.1 Объемы различных форм учебной работы и виды контроля в соответствии с учебным планом (заочная форма)**

Общая трудоемкость дисциплины составляет   3   зачетных единицы.

Виды учебной работы	Всего ак. ч	Распределение трудоемкости по семестрам, ак. ч
		7 семестр
Общая трудоемкость дисциплины	108	108
<b>Контактная работа, в т.ч. аудиторные занятия:</b>	9,5	9,5
Лекции	4	4
<i>в т.ч. в форме практической подготовки</i>	-	-
Практические занятия (ПЗ)	4	4
<i>в т.ч. в форме практической подготовки</i>	4	4
Консультации текущие	0,6	0,6
Контрольная работа	0,8	0,8
Виды аттестации (зачет, экзамен)	0,1	0,1
<b>Самостоятельная работа:</b>	94,6	94,6
Проработка материалов по конспекту лекций и учебникам (собеседование, тестирование)	64,6	64,6
Выполнение расчетов для практических работ	15	15
Оформление отчетов по практическим работам	15	15
Контроль	3,9	3,9

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ  
ДЛЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ**

по дисциплине

**Прикладная статистика**

## 1 Перечень компетенций с указанием этапов их формирования

№ п/п	Код компетенции	Формулировка компетенции	Код и наименование индикатора достижения компетенции
1	УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	ИД1 <sub>УК-1</sub> Умеет анализировать и систематизировать разнородные данные, оценивать эффективность процедур анализа проблем и принятия решений в профессиональной деятельности.
2	УК-2	Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений	ИД1 <sub>УК-2</sub> Умеет анализировать альтернативные варианты решений для достижения намеченных результатов; разрабатывать план, определять целевые этапы и основные направления работ.
			ИД2 <sub>УК-2</sub> Владеет методиками разработки цели и задач проекта; методами оценки продолжительности и стоимости проекта, а также потребности в ресурсах.
3	ПКв-4	Способность составлять технико-экономическое обоснование проектных решений и техническое задание на разработку информационной системы	ИД1 <sub>ПКв-4</sub> Умеет составлять технико-экономического обоснования проектных решений и техническое задание на разработку ИС
4	ПКв-7	Способность настраивать, эксплуатировать и сопровождать информационные системы и сервисы	ИД1 <sub>ПКв-7</sub> Участвует в экспертном тестировании ИС на этапе опытной эксплуатации

## 2 Паспорт оценочных материалов по дисциплине

№ п/п	Разделы дисциплины	Индекс контролируемой компетенции (или ее части)	Оценочные материалы		Технология/процедура оценивания (способ контроля)
			наименование	№№ заданий	
5 семестр					
1	Выбор способов решения задач в рамках поставленной цели. Язык R	УК-1 УК-2	Банк тестовых заданий	1-21	Бланочное или компьютерное тестирование
			Вопросы к зачету	111-124	Проверка преподавателем
			Кейс-задания	59-73	Проверка преподавателем
2	Проверка статистических гипотез как оптимальный способ решения для достижения намеченных результатов	УК-2 ПКв-4	Банк тестовых заданий	22-38	Бланочное или компьютерное тестирование
			Вопросы к зачету	125-130	Проверка преподавателем
			Кейс-задания	74-88	Проверка преподавателем
3	Регрессионный анализ данных с целью оценки эффективности принятия решений в профессиональной деятельности	УК-1 ПКв-4	Банк тестовых заданий	39-50	Бланочное или компьютерное тестирование
			Кейс-задания	89-103	Проверка преподавателем
			Вопросы к практическим работам	131-138	Проверка преподавателем
4	Методы отбора признаков как один	УК-1 ПКв-7	Банк тестовых заданий	51-58	Бланочное или компьютерное

	из вариантов решения задачи для достижения намеченных результатов				тестирование
			Вопросы к практическим работам	139-145	Проверка преподавателем
			Кейс-задания	104-109	Проверка преподавателем
5	Использование методов предиктивной аналитики при составлении технико-экономического обоснования проектных решений	УК-1 УК-2 ПКв-4 ПКв-7	РПР	110	Проверка преподавателем
			Вопросы к практическим работам	146-149	Проверка преподавателем

### 3 Оценочные средства для промежуточной аттестации

**Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы**

Аттестация обучающегося по дисциплине проводится в форме тестирования или письменного ответа, выполнения расчетно-практической работы и предусматривает возможность последующего собеседования (зачета).

Каждый вариант теста включает 30 контрольных заданий на проверку знаний.

Каждый билет включает 3 контрольных вопросов, из них:

- 2 контрольных вопроса на проверку знаний;
- 1 контрольный вопрос на проверку умений и навыков.

#### 3.1. Тесты (тестовые задания)

№ задания	Текст задания
1	Статистические пакеты по признаку функциональности могут быть разделены на ____ основных групп: 2; <b>3;</b> 4; 5;
2	Перечислите минимальный набор статистических методов анализа, который включен во все пакеты: <b>описательная статистика;</b> <b>дисперсионный анализ;</b> <b>кластерный анализ;</b> <b>обработка данных;</b>
3	Какие требования предъявляются к статистическим пакетам общего назначения? соответствие характеру решаемых задач; <b>объем обрабатываемых данных;</b> <b>требования, предъявляемые к квалификации пользователя (уровень знаний в области статистики);</b> <b>имеющееся в наличии компьютерное оборудование;</b>
4	К какой группе пакетов относится пакет SPSS? <b>универсальный;</b> профессиональный; специализированный;
5	К какой группе пакетов относится пакет Stadia? <b>универсальный;</b> профессиональный;

	специализированный;
6	К какой группе пакетов относится пакет Minitab? <b>универсальный;</b> профессиональный; специализированный;
7	К какой группе пакетов относится пакет SAS? универсальный; <b>профессиональный;</b> специализированный;
8	К какой группе пакетов относится пакет BMDP? универсальный; <b>профессиональный;</b> специализированный;
9	К какой группе пакетов относится пакет BioStat? универсальный; профессиональный; <b>специализированный;</b>
10	К какой группе пакетов относится пакет Datascope? универсальный; профессиональный; <b>специализированный;</b>
11	С какого сайта следует скачать последнюю версию R? <a href="https://www.r-project.org/">https://www.r-project.org/</a> ; <a href="https://www.project-r.org/">https://www.project-r.org/</a> . <a href="https://www.r-language.org/">https://www.r-language.org/</a> . <a href="https://www.rlanguage.org/">https://www.rlanguage.org/</a> .
12	Как проверить версию установленного языка R? <b>'R.version.string'</b> 'R.version??' 'R??' '?R.version?'
13	Какие типы данных предусмотрены в R? <b>numeric;</b> real; main; <b>logical;</b>
14	Какая функция осуществляет переход к факторным переменным? <b>Factor;</b> Fact; Factorial; Fac;
15	Можно ли в R задать матрицу через функцию array()? <b>Да;</b> Нет;
16	Каким атрибутом массив отличается от вектора? <b>Dim;</b> Div; Den; Day;
17	Как задать класс типа СПИСОК? Spis; <b>List;</b> Make_list; Do_list;
18	Как представить данные в табличном виде? Table; Data.set; <b>Data.frame;</b> Table.set;
19	Какая функция выводит информацию о структуре данных? Head;

	Body; <b>Str;</b> InfDate;
20	Какая функция отвечает за построение графиков в R? Plot; Paint; Graph; <b>Ggplot;</b>
21	Каким образом отобразить на графике плотность распределения? <b>Density;</b> Plotn_rasp; Graph_dens; Adjust;
22	Какая буква при работе с распределениями в R отвечает за функцию плотности вероятности? <b>D;</b> P; Q; R;
23	Какая буква при работе с распределениями в R отвечает за кумулятивную функцию распределения вероятности? D; <b>P;</b> Q; R;
24	Какая буква при работе с распределениями в R отвечает за функцию для нахождения квантилей? D; P; <b>Q;</b> R;
25	Какая буква при работе с распределениями в R отвечает за функцию генерации случайных чисел? D; P; Q; <b>R;</b>
26	Каким образом в языке R записывается биномиальное распределение? Bin_rasp; <b>Binom;</b> Binomial; RaspBin;
27	Какое распределение в языке R записывается как unif? Распределение Коши; Гамма-распределение; <b>Равномерное распределение;</b> Полиномиальное распределение;
28	Какое распределение в языке R записывается как multinom? Распределение Коши; Гамма-распределение; Равномерное распределение; <b>Полиномиальное распределение;</b>
29	Сколько типов кривых Пирсона принято различать на практике? 4; 5; 6; <b>7;</b>
30	Относится ли быстрый (грубый) критерий к числу непараметрических? <b>Да;</b> Нет;
31	Каково другое название критерия Манна-Уитни-Вилкоксона? Критерий сдвига вправо;

	Критерий сдвига влево; <b>Критерий сдвига среднего;</b> Критерий правого сдвига;
32	Каково другое название критерия Клотца? Критерий изменения среднего; Критерий изменения матожидания; <b>Критерий изменения дисперсии;</b> Критерий изменения отклонения;
33	Каково другое название критерия Ансари-Брэдли? Критерий изменения среднего; Критерий изменения матожидания; <b>Критерий изменения дисперсии;</b> Критерий изменения отклонения;
34	Относится ли критерий Сатервайта к числу непараметрических? Да; <b>Нет;</b>
35	Относится ли критерий Уэлча к числу параметрических? <b>Да;</b> Нет;
36	Какой параметрический критерий применяется в случае нормального распределения? Критерий Ивановского; Критерий Петровского; <b>Критерий Романовского;</b> Критерий Федоровского;
37	Критерий Фишера относится к параметрическим критериям в случае _____ распределения. <b>Экспоненциального;</b> Нормального; <b>Биномиального;</b> Логарифмического;
38	Какой пакет языка R надо использовать для определения соответствия данных «под распределение»? Distrplus; <b>Fitdistrplus;</b> Raspdisp; Fitrasp;
39	Статистической гипотезой называется: сравнение двух исследуемых оценок, основанных на соотношении их дисперсий; <b>предположение о свойстве генеральной совокупности, которое можно проверить, опираясь на данные выборки;</b> предположение о свойстве состоятельности, которое является одним из асимптотическим свойством; гипотеза о распределении и структурных характеристиках.
40	Параметрической гипотезой называется: <b>гипотеза о параметрах генеральной совокупности;</b> гипотеза о параметрах состоятельности; гипотеза о параметрах эффективности; гипотеза о параметрах несостоятельности.
41	Нулевой гипотезой называется: <b>гипотеза о том, что две совокупности, сравниваемые по одному или нескольким признакам, не отличаются;</b> гипотеза о том, что две состоятельности, сравниваемые по одному или нескольким признакам, не отличаются; гипотеза о том, что две генеральные состоятельности, сравниваемые по одному или нескольким признакам, не отличаются; гипотеза о том, что две несостоятельности, сравниваемые по одному или нескольким признакам, не отличаются.
42	Статистическим критерием называется: <b>определенное правило, устанавливающее условия, при которых проверяемую нуле-</b>



	<p><b>вую гипотезу следует либо отклонить, либо не отклонять;</b>  критерий, позволяющий сравнивать одну или несколько гипотез, которые отличаются по некоторым признакам;  критерий, отвергающий результат, который при истинности выдвинутой нулевой гипотезы маловероятен;  критерий, формирующийся в виде статистической гипотезы.</p>
43	<p>Уравнение, описывающее корреляционную связь между парой переменных: зависимой переменной (результатом) <math>y</math> и независимой переменной (фактором) <math>x</math>, называется:  <b>парная регрессия.</b></p>
44	<p>Уравнение, отражающее корреляционную связь между результатом и несколькими факторами, называется:  <b>модель множественной регрессии;</b>  <b>модель суммарной регрессии;</b>  <b>модель производной регрессии.</b></p>
45	<p>Под методом последовательного включения факторов понимается:  построение модели с максимально большим количеством факторов, из которой затем поочередно исключаются незначимые факторы до тех пор, пока модель не будет иметь только значимые параметры при факторах;  построение парного уравнения регрессии с наиболее значимым (по степени влияния на результат) фактором; нахождение случайных остатков; построение уравнения зависимости от следующего по степени влияния на результат фактора;  <b>построение модели с фактором, наиболее тесно связанным с результатом; поочередное добавление других факторов; оценка целесообразности включения нового фактора с точки зрения сокращения остаточной дисперсии;</b>  построение модели начинается с расчета парной регрессии с фактором, наиболее тесно связанным с результатом; добавление каждого нового фактора сопровождается не только оценкой значимости включения данного фактора, но и проверкой значимости влияния на результат прочих факторов.</p>
46	<p>Под методом исключения факторов понимается:  <b>построение модели с максимально большим количеством факторов, из которой затем поочередно исключаются незначимые факторы до тех пор, пока модель не будет иметь только значимые параметры при факторах;</b>  построение парного уравнения регрессии с наиболее значимым (по степени влияния на результат) фактором; нахождение случайных остатков; построение уравнения зависимости от следующего по степени влияния на результат фактора;  построение модели с фактором, наиболее тесно связанным с результатом; поочередное добавление других факторов; оценка целесообразности включения нового фактора с точки зрения сокращения остаточной дисперсии;  построение модели начинается с расчета парной регрессии с фактором, наиболее тесно связанным с результатом; добавление каждого нового фактора сопровождается не только оценкой значимости включения данного фактора, но и проверкой значимости влияния на результат прочих факторов.</p>
47	<p>Под шаговым регрессивным анализом понимается:  построение модели с максимально большим количеством факторов, из которой затем поочередно исключаются незначимые факторы до тех пор, пока модель не будет иметь только значимые параметры при факторах;  построение парного уравнения регрессии с наиболее значимым (по степени влияния на результат) фактором; нахождение случайных остатков; построение уравнения зависимости от следующего по степени влияния на результат фактора;  построение модели с фактором, наиболее тесно связанным с результатом; поочередное добавление других факторов; оценка целесообразности включения нового фактора с точки зрения сокращения остаточной дисперсии;  <b>построение модели начинается с расчета парной регрессии с фактором, наиболее</b></p>

	<b>тесно связанным с результатом; добавление каждого нового фактора сопровождается не только оценкой значимости включения данного фактора, но и проверкой значимости влияния на результат прочих факторов.</b>
48	Под ступенчатым регрессивным анализом понимается: построение модели с максимально большим количеством факторов, из которой затем поочередно исключаются незначимые факторы до тех пор, пока модель не будет иметь только значимые параметры при факторах; <b>построение парного уравнения регрессии с наиболее значимым (по степени влияния на результат) фактором; нахождение случайных остатков; построение уравнения зависимости от следующего по степени влияния на результат фактора;</b> построение модели с фактором, наиболее тесно связанным с результатом; поочередное добавление других факторов; оценка целесообразности включения нового фактора с точки зрения сокращения остаточной дисперсии; построение модели начинается с расчета парной регрессии с фактором, наиболее тесно связанным с результатом; добавление каждого нового фактора сопровождается не только оценкой значимости включения данного фактора, но и проверкой значимости влияния на результат прочих факторов.
49	Точечный прогноз – это: прогнозируемый расчет, на основе графического исчисления данных по уравнению парной регрессии; значение результата, полученное графически с помощью уравнения множественной регрессии; предполагаемое значение результата для прогнозирования значений факторов; <b>расчетное значение результата, полученное подстановкой в уравнение множественной регрессии прогнозных значений факторов.</b>
50	Интервальный прогноз – это: минимальное значение результата, с заданной долей вероятности попадает в фактическое значение результата при заданных прогнозных значениях факторов; максимальное значение результата, с заданной долей вероятности попадает в фактическое значение результата при заданных прогнозных значениях факторов; <b>минимальное и максимальное значение результата, в промежутке между которыми с заданной долей вероятности попадает фактическое значение результата при заданных прогнозных значениях факторов;</b> верного определения нет.
51	Какой подход в анализе регрессии по всем переменным является первым? Регрессия по основным подмножествам; <b>Регрессия по всем подмножествам;</b> Регрессия по дополнительным подмножествам; Регрессия по первому подмножеству;
52	Как определить оптимальную модель из множества регрессий по всем подмножествам? По остаточной ошибке; По остаточной дисперсии; <b>По коэффициенту детерминации;</b> По математическом ожиданию;
53	Как определить оптимальную модель из множества регрессий по всем подмножествам? По остаточной ошибке; По остаточной дисперсии; <b>По величине байесовского информационного критерия;</b> По математическом ожиданию;
54	Как по величине байесовского информационного критерия (BIC) определить, какая модель лучше? Чем выше BIC, тем лучше; <b>Чем ниже BIC, тем лучше;</b> Чем ближе BIC к 0, тем лучше; Чем ближе BIC по модулю к 1, тем лучше;
55	В чем заключается суть step-wise регрессии? Строится модель регрессии по всем переменным, а потом происходит пошаговое удаление переменных;

	<p>Строится модель регрессии по всем переменным, а потом происходит пошаговое прибавление переменных;</p> <p>Строится модель регрессии по одной переменной, а потом происходит пошаговое добавление переменных;</p> <p>Строится модель регрессии по двум переменным, а потом происходит пошаговое добавление переменных;</p>
56	<p>В чем состоит суть метода случайного леса для отбора признаков?</p> <p><b>Он выдает числовые значения «важности» каждой переменной уравнения регрессии;</b></p> <p>Он выдает оптимальное количество переменной уравнения регрессии;</p> <p>Он выдает переменные уравнения регрессии в порядке их следования в модели;</p> <p>Он выдает числовые значения проверки адекватности модели;</p>
57	<p>Какая функция в языке R отвечает за регрессию с регуляризацией?</p> <p>Regsubset;</p> <p>Cforest;</p> <p><b>Glmnet;</b></p> <p>Relimportance;</p>
58	<p>Какой пакет языка R позволяет оценить вклад каждой переменной в модель?</p> <p>Glmnet;</p> <p>Party;</p> <p>MASS;</p> <p><b>Relaimpo;</b></p>

### 3.2. Кейс-задания

№ задания	Текст задания
59	<p>Задать вектор A, состоящий из элементов от 5 до 10</p> <p><b>Решение:</b> A(5,6,7,8,9,10)</p>
60	<p>Задать вектор B, состоящий из элементов от 5 до 15 с шагом 2</p> <p><b>Решение:</b> B&lt;-seq(from=5, to=15, by=2)</p>
61	<p>Задать вектор C, состоящий из элементов от 5 до 25</p> <p><b>Решение:</b> C&lt;-5:25</p>
62	<p>Задать вектор D, состоящий из 16 элементов 7</p> <p><b>Решение:</b> D&lt;-rep(7,16)</p>
63	<p>Задать вектор E, состоящий из 6 комбинаций элементов 7,5,3</p> <p><b>Решение:</b> E&lt;-rep(c(7,5,3),6)</p>
64	<p>Задать вектор F, состоящий из 4 цифр 11, 7 и 2</p> <p><b>Решение:</b> F&lt;-rep(c(11,7,2), each=4)</p>
65	<p>Определить длину вектора F.</p> <p><b>Решение:</b> Length (F)</p>
66	<p>Определить последний элемент вектора E.</p> <p><b>Решение:</b> E[Length (E)]</p>
67	<p>Записать с третьего по седьмой элементы вектора D.</p> <p><b>Решение:</b> D[3:7]</p>
68	<p>Определить порядковые номера каждого числа в векторе C.</p> <p><b>Решение:</b> Order (C)</p>
69	<p>Сформировать из исходного вектора C вектор, упорядоченный по возрастанию.</p> <p><b>Решение:</b> C[Order (C)]</p>

70	Сформировать из исходного вектора C вектор, упорядоченный по убыванию. <b>Решение:</b> C[Order (C, decreasing= TRUE)]
71	Добавить в исходный вектор B число 55 на четвертую позицию <b>Решение:</b> B<-c(B[1:3],55,B[4:10])
72	Определить остатки от деления всех элементов вектора C на 5. <b>Решение:</b> C%%5
73	Убрать 4-й элемент вектора A. <b>Решение:</b> A[-4]
74	Убрать со 2-го по 5-й элементы вектора B. <b>Решение:</b> B[-2:-5]
75	Проверить все элементы вектора C на равенство числу 12. <b>Решение:</b> C==12
76	Определить количество элементов вектора C, равных числу 8. <b>Решение:</b> Sum(C==8)
77	Проверить все элементы вектора D на то, что они все больше числа 5. <b>Решение:</b> All(D>5)
78	Проверить все элементы вектора D на то, что хотя бы одно из них больше числа 5. <b>Решение:</b> Any(D>5)
79	Проверить все элементы вектора A на то, что они больше 5 и меньше 13. <b>Решение:</b> A>5 & A<13
80	Выбрать все элементы вектора B, которые больше 3 и меньше 9. <b>Решение:</b> B [B>3 & B<9]
81	Выбрать все элементы вектора E, которые при делении на три дают в остатке 1 или квадрат которых больше 50. <b>Решение:</b> E[E%%3==1   E*E>50]
82	Определить индексы элементов вектора F, куб которых меньше 77 <b>Решение:</b> Which (F*F*F<77)
83	Из векторов B и C получить новый вектор, содержащий неповторяющиеся элементы B и C <b>Решение:</b> Unique (B+C)
84	Как запустить на выполнение датафрейм с названием stud <b>Решение:</b> df<-stud
85	Каким образом узнать, есть ли пропущенные наблюдения в базе данных df <b>Решение:</b> any(is.na(df)==TRUE)
86	Каким образом исключить из базы данных df пропущенные наблюдения <b>Решение:</b> df<-na.omit(df)
87	Каким образом можно посмотреть базу данных df <b>Решение:</b> Glimpse(df)
88	Каким образом можно построить уравнение линейной регрессии между переменными Age и Stat базы данных stud <b>Решение:</b> Model<-lm(data = stud, age ~ stat)
89	Каким образом можно посмотреть результаты расчетов по модели линейной регрессии model

Решение:  
Summary (Model)

### 3.3. Расчетно-практическая работа

№ задания	Текст вопроса
90	<p>На основе анализа графиков исходного ряда, его первых и вторых разностей сделать предположение о типе процесса, представленного временным рядом. На основе анализа коррелограмм АКФ и ЧАКФ сделать предположение о типе процесса, представленного временным рядом. Коррелограммы анализируются как для исходного ряда, так и для ряда первых и вторых разностей. С помощью расширенного теста Дики-Фуллера провести идентификацию случайного процесса, применяя процедуру Доладо-Дженкенсона-Сосвила-Риверо. Сделать заключение о типе процесса.</p> <p>Проанализировать ряд на наличие структурных изломов и скачков с помощью команды <code>breakpoints</code> пакета <code>strucchange</code>:</p> <ul style="list-style-type: none"><li>-выбрать точку излома и для неё создать фиктивные переменные <math>(ds, ds1, dt)</math>, отвечающие за тип структурного скачка (дополнительно можно перебрать структурные изменения для всех точек);</li><li>-в зависимости от типа процесса построить модели с включением соответствующих фиктивных переменных, отвечающих за тип структурного скачка.</li></ul> <p>-проверить статическую значимость коэффициентов при фиктивных переменных с помощью теста Стьюдента, а также общую значимость уравнения для теста, проверить значимость статистики Дики-Фуллера (коэффициент при лаговой переменной);</p> <p>-сделать вывод о типе процесса с учетом структурных изменений</p> <p>Для построения и мониторинга моделей ARIMA/SARIMA:</p> <ul style="list-style-type: none"><li>-перевести временной ряд, измеряющий исследуемый показатель в стационарный, в зависимости от того, к какому типу процессов относится моделируемый ряд;</li><li>-на основе анализа коррелограмм АКФ и ЧАКФ, полученного стационарного ряда провести идентификацию порядков моделей ARIMA/SARIMA;</li><li>-методом наименьших квадратов найти оценки параметров моделей ARIMA/SARIMA;</li><li>-проведение процедуры адекватности, отобранной модели, сделать выводы о статистической значимости коэффициентов модели, о показателях качества модели;</li><li>-проанализировать коррелограммы АКФ и ЧАКФ остатков модели для определения их соответствия белому шуму. Проанализировать Q-Q график остатков на соответствие их нормальному закону распределению;</li><li>-при построении нескольких статистически значимых моделей ARIMA/SARIMA провести процедуру их селекции, исходя из минимума информационных критериев.</li></ul> <p>Оценить спецификацию модели ARIMAX с учётом определенных структурных изломов/скачков/подскоков, наличия или отсутствия тренда и определенных ранее на основе анализа коррелограмм порядков модели авторегрессии <math>p</math>, скользящего среднего <math>q</math>, сезонной авторегрессии <math>P</math>, сезонной скользящей средней <math>Q</math>. Для определенной точки излома создать фиктивные переменные <math>(ds, ds1, dt)</math>, отвечающие за тип структурного скачка, а также переменную тренда. Создать датафрейм для созданных 4-х переменных. Оценить модель как ARIMAX с введением экзогенных переменных (отвечающих за структурные изменения и переменную тренда). Проверить статическую значимость коэффициентов при фиктивных переменных, отвечающих за структурные изменения, и коэффициентов SARIMA помощью теста Стьюдента. Оценить соответствие остатков белому шуму с помощью коррелограммы АКФ и соответствие нормальному распределению с помощью QQ-графика. Сформировать отчет и сделать вывод о модели ARIMAX с учетом структурных изменений.</p> <p>Построить модель Брауна, подобрав параметр адаптации в автоматическом режиме. Определить параметр адаптации. Проанализировать качество модели на основе анализа соответствия остатков белому шуму с помощью коррелограмм, и анализа QQ-графика на соответствие остатков нормальному распределению. Построить модель Хольта, подобрав параметры адаптации в автоматическом режиме. Определить параметры адаптации. Проанализировать качество модели на основе анализа соответствия остатков белому шуму с помощью коррелограмм, и анализа QQ-графика на соответствие остатков нормальному распределению. Построить аддитивную модель Хольта-Уинтерса, подобрав параметр адаптации в автоматическом режиме. Опреде-</p>

	<p>лить параметры адаптации. Проанализировать качество модели на основе анализа соответствия остатков белому шуму с помощью коррелограмм, и анализа QQ-графика на соответствие остатков нормальному распределению. Определить параметры адаптации. Построить мультипликативную модель Хольта-Уинтерса, подобрав параметры адаптации в автоматическом режиме. Проанализировать качество модели на основе анализа соответствия остатков белому шуму с помощью коррелограмм, и анализа QQ-графика на соответствие остатков нормальному распределению. Сравнить метрики прогнозных качеств модели, сравнить значения информационных критериев. Выбрать наиболее подходящую для прогнозирования модель адаптивного сглаживания.</p>
--	--

### 3.4. Вопросы к зачету

№ вопроса	Формулировка вопроса
111	<p>Требования к статистическим пакетам обработки данных</p> <p><b>Ответ</b></p> <p>Перед пользователями различных категорий встает вопрос выбора оптимального статистического пакета для поиска верных ответов на существующие вопросы. Очевидно, что оптимальным является вариант, сочетающий в себе необходимые функциональные возможности, высокое качество работы и умеренную цену. При выборе пакета учитываются следующие параметры:</p> <ul style="list-style-type: none"> <li>соответствие характеру решаемых задач;</li> <li>объем обрабатываемых данных;</li> <li>требования, предъявляемые к квалификации пользователя (уровень знаний в области статистики);</li> <li>имеющееся в наличии компьютерное оборудование.</li> </ul>
112	<p>Группы статистических пакетов</p> <p><b>Ответ</b></p> <p>Статистические пакеты по признаку функциональности могут быть разделены на 3 основные группы.</p> <p><b>Универсальные пакеты</b>, или пакеты общего назначения. Эти пакеты не ориентированы на специфическую предметную область и могут применяться для анализа данных из различных областей деятельности. Как правило, они предлагают широкий диапазон статистических методов и имеют относительно простой интерфейс.</p> <p><b>Профессиональные пакеты</b> (например, SAS, BMDP).</p> <p>Профессиональные пакеты отличаются от универсальных тем, что позволяют работать со сверхбольшими объемами данных, применять узкоспециализированные методы анализа, создавать собственную систему обработки данных.</p> <p><b>Специализированные пакеты</b> В некоторых областях деятельности анализируемые данные настолько специфичны, что к ним следует применять особые методы статистического анализа, как правило, не представленные в универсальных пакетах.</p> <p>Специализированные пакеты позволяют проводить анализ с использованием ограниченного числа специализированных статистических методов или применимы к использованию для решения вопросов, относящихся к отдельно взятой предметной области. Как правило, с подобными статистическими пакетами работают специалисты, хорошо знакомые с методами анализа данных в той области, на которую ориентирован пакет.</p>

113	<p><b>Обзор статистического пакета SAS</b></p> <p><b>Ответ</b></p> <p>Одним из популярных статистических приложений является система SAS, которая зародилась в 60-х годах XX века в Университете Северной Каролины как приложение для анализа результатов сельскохозяйственных исследований. На сегодняшний день система продолжает развиваться в компании SAS Institute, которая выпустила уже девятую версию этой программы. Область применения SAS - самые разнообразные научные исследования, бизнес аналитика и т. д.</p> <p>Система состоит из модулей, каждый из которых выполняет определенный круг задач. Наиболее часто при <u>статобработке</u> используются модули BASE и STAT. В системе SAS реализован собственный язык программирования, который по своему синтаксису ближе к Basic и не похож на R или S. Система позволяет загружать данные из внешних файлов или же вводить их непосредственно в окно терминала. Работая с использованием SAS можно проводить статистическую обработку данных разного уровня сложности, в соответствии с поставленными задачами. Взаимодействие с программой возможно как в консольном режиме, так и через графический интерфейс, который представляет собой графическую оболочку для упрощенного ввода команд языка программирования SAS.</p> <p>Если говорить о компонентах SAS, то в SAS доступно более 200 компонентов.</p> <p>В основном используется <u>Window SAS</u> как в организации, так и в учебных заведениях. Некоторые организации используют <u>Linux</u>, но нет графического интерфейса пользователя, поэтому вы должны писать код для каждого запроса. Но в окне SAS есть много доступных утилит, которые очень помогают программистам, а также сокращают время написания кодов.</p> <p>Окно <u>SaS</u> состоит из 5 частей.</p> <p>Пакет предоставляет пользователям следующие возможности статистического анализа данных:</p> <p>смешанные модели; анализ <u>категорийных</u> данных; байесовский анализ; многомерный анализ; анализ выживаемости; психометрический анализ; анализ пропущенных значений.</p> <p>Достоинства SAS:</p> <p>быстрая обработка очень больших объемов данных; возможность преобразования математических формул в программный код; создание пользовательских модулей; получение консультативной помощи в выборе методов анализа и в интерпретации его результатов, а также рекомендаций по дальнейшей работе с исходными данными; совместимость с операционными системами <u>Linux</u> и <u>Windows</u>.</p> <p>Недостатки SAS:</p> <p>сложен в освоении для неспециалистов; высокая стоимость</p>
114	<p><b>Обзор статистического пакета Stata</b></p> <p><b>Ответ</b></p>

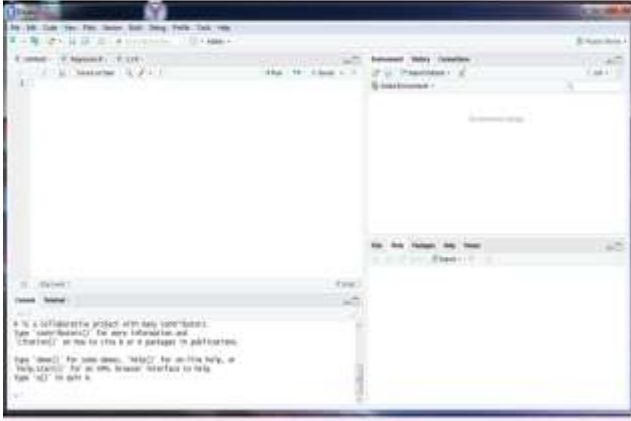
	<p>Пакет <i>Stata</i> – универсальный статистический пакет компании <i>StataCorp</i><sup>7</sup>. Пакет <i>Stata</i> позиционируется как инструмент анализа, предназначенный для специалистов, которые занимаются научными исследованиями. <i>Stata</i> предоставляет пользователям следующие возможности статистического анализа данных:</p> <ul style="list-style-type: none"> <li>-динамические факторные модели, линейные, обобщенные линейные модели и нелинейные модели, многоуровневые смешанные модели, модели выживания;</li> <li>-обобщенные оценки уравнений (GEE);</li> <li>-непараметрические методы, методы повторных выборок и статистического эксперимента;</li> <li>-проверка гипотез и работа с оцененными моделями;</li> <li>-максимизация функций правдоподобия, заданных пользователем;</li> <li>-матричные команды.</li> </ul> <p>Предлагаются четыре различных решения при установке <i>Stata</i>: <i>Stata/IC</i>, <i>Stata/SE</i>, <i>Stata/MP</i>, <i>Small Stata</i>.</p> <p>Достоинства <i>Stata</i>:</p> <ul style="list-style-type: none"> <li>-широкий набор средств статистического анализа данных;</li> <li>-широкий спектр возможностей по управлению данными (возможности комбинирования и изменения наборов данных, управления переменными и пр.);</li> <li>-возможность программирования всей последовательности команд (начиная от загрузки данных до проведения всех этапов их анализа);</li> <li>-использование координатного указателя;</li> <li>-интерактивный справочник;</li> <li>-создание графики полиграфического качества;</li> <li>-возможность написания собственных программных модулей;</li> </ul> <p>Совместимость с операционными системами <i>Windows</i>, <i>Macintosh</i> и <i>Unix</i> (включая <i>Linux</i>);</p> <ul style="list-style-type: none"> <li>-полное документационное сопровождение ("Базовое справочное руководство" с более чем 1800 страницами информации, содержащее подробные примеры и ссылки на литературу по статистике);</li> <li>-экспорт результатов в <i>MS Office</i> и <i>SAS</i>;</li> <li>-для работы программы требуется лишь 512 MB оперативной памяти, сама программа занимает 250 MB на жестком диске компьютера;</li> <li>-возможность запуска команды, как через соответствующий пункт меню, так и через написание команды в командной строке.</li> </ul> <p>Недостатки <i>Stata</i>:</p> <ul style="list-style-type: none"> <li>-отсутствие возможности полноценного экспорта и импорта данных в базы данных, электронные таблицы и текстовые процессоры;</li> <li>-возможность прямого открытия в программе лишь файлов с разрешением <i>.dta</i>.</li> </ul>
115	<p><b>Обзор статистического пакета MedCalc</b></p> <p><b>Ответ</b></p> <p>Программа <i>MedCalc</i>, разрабатываемая с 1993 года бельгийской компанией <i>MedCalc Software</i>, позиционируется как полноценное статистическое приложение, созданное в соответствии с потребностями исследователей биомедицинского направления. Разработчики акцентируют внимание исследователей на удобстве использования <i>MedCalc</i> для анализа ROC-кривых. Программа удобна тем, что не предлагает избыточной функциональности, которая зачастую приводит в замешательство неподготовленного человека, начинающего работать с универсальными приложениями. В дополнение к этому, возможность работы только в графическом интерфейсе без использования командной строки делает программу менее гибкой, но зато более привлекательной для использования в данной области науки, поскольку специалисты с медицинским образованием очень редко могут похвастаться большим опытом работы с математическими программами.</p> <p>На сегодняшний день создана уже двенадцатая версия программы. Работать в <i>MedCalc</i> могут только пользователи <i>Windows</i>, но этот недостаток компенсируется относительно малыми системными требованиями и возможностью запустить приложение как в <i>Windows 2000</i>, так и в <i>Windows 7</i>. Для тех, кто еще ни разу не пользовался программой, есть возможность скачать полнофункциональную демонстрационную версию продукта с сайта <i>medcalc.org</i>, которая будет работать без ограничений в течение пятнадцати дней. Кроме того, в комплект поставки входят демонстрационные файлы, содержащие наборы данных и примеры их анализа.</p> <p>Ввод данных в <i>MedCalc</i> осуществляется в интегрированном табличном редакторе или же путем импорта файлов различных форматов, таких как <i>*.csv</i>, <i>excel</i> и др. Для вызова встроенного редактора достаточно выбрать команду <i>Spreadsheet</i> в меню, после чего можно приступать к формированию таблицы. В статистических программах колонки таблиц называют «переменными», а строки «случаями». При создании таблицы полезным будет соблюдение нескольких правил:</p> <ul style="list-style-type: none"> <li>• Первая переменная должна содержать порядковые номера случаев. Это необходимо для того, чтобы после пересортировки значений иметь возможность восстановить их прежний порядок.</li> <li>• Числовые значения следует вводить без округления, чтобы не терять информацию.</li> <li>• При отсутствии некоторых значений их можно пропустить, оставив в таблице пустые ячейки.</li> <li>• В каждой переменной для каждого случая должно быть только одно значение.</li> </ul>
116	<p><b>Обзор статистического пакета SPSS Statistics</b></p> <p><b>Ответ</b></p>



	<p>Пакет SPSS (Statistical Package for the Social Sciences) – универсальный статистический пакет компании SPSS Inc. Первая версия пакета была выпущена в 1968 г. В 2009 г. компания IBM поглотила SPSS Inc., поэтому новая версия пакета включает в свое название аббревиатуру IBM (IBM SPSS Statistics 19). SPSS является модульной программой. Ее основу составляет базовый модуль (SPSS Base), позволяющий осуществлять управление данными и содержащий наиболее распространенные методы статистического анализа данных: проведение описательной статистики; построение линейных и нелинейных моделей; осуществление преобразования данных; проведение факторного, кластерного, дисперсионного анализов; вычисление корреляций; построение графиков; подготовка отчетов и пр.</p> <p>Для проведения расширенного и углубленного анализа данных могут быть установлены дополнительные модули пакета. Для пакета IBM SPSS Statistics 19 разработаны 16 различных модулей.</p> <p>Достоинства SPSS:</p> <ul style="list-style-type: none"> <li>развитый аппарат статистического анализа;</li> <li>универсальность (может быть использован для решения широкого круга вопросов из различных предметных областей, требующих проведения статистического анализа данных);</li> <li>широкий набор статистических и графических процедур (более 50 типов диаграммы) анализа данных, а также процедур создания отчетов;</li> <li>высокая скорость вычислений, простой и удобный интерфейс;</li> <li>детальная контекстно-ориентированная справочная система, позволяющая неопытному пользователю с большей легкостью ориентироваться в программе;</li> <li>возможность свободного скачивания демонстрационной версии продукта на официальном сайте компании, наличие версий продукта на различных языках;</li> <li>совместимость с операционными системами Windows, Mac, Linux;</li> <li>наличие значительного количества литературы по работе с пакетом.</li> </ul> <p>Недостатки SPSS:</p> <ul style="list-style-type: none"> <li>высокие требования к системе компьютера (требуется 1GB оперативной памяти, 800MB памяти на жестком диске и процессор с частотой 1GHz и выше);</li> <li>высокая цена по сравнению со статистическими пакетами аналогичного уровня (стоимость покупки для индивидуального пользования сроком на год составляет около 1000 долл.б).</li> </ul> <p>Последняя версия SPSS включает в себя следующие новые возможности:</p> <ul style="list-style-type: none"> <li>импорт данных из Excel и SAS;</li> <li>экспорт результатов в MS Office, PDF; сохранение результатов в формате HTML;</li> <li>одновременная работа с несколькими наборами данных;</li> <li>построение диаграммы для переменных с множественными ответами;</li> <li>построение диаграммы с двумя осями Y;</li> <li>улучшенный редактор синтаксиса с поддержкой автозавершения и цветового кодирования команд;</li> <li>быстрая подготовка данных к анализу посредством Автоматизированной подготовки данных (IBM SPSS Data Preparation), позволяющей облегчить процесс интеллектуального анализа данных, выявляя и исправляя ошибки в данных и объясняя пропущенные значения. Также посредством этой функции можно подготовить отчет с рекомендациями о возможности использования данных для анализа.</li> </ul>
117	<p>Обзор статистического пакета Statistica</p> <p><b>Ответ</b></p>

	<p>Пакет STATISTICA – универсальный статистический пакет компании StatSoft Inc1. Первая версия пакета (STATISTICA for DOS) была выпущена в 1991 г. На сегодняшний день разработана 9-я версия пакета (STATISTICA 9). Русифицирована только 6-я версия (STATISTICA 6). Стандартную комплектацию пакета STATISTICA составляют три модуля, которые могут приобретаться как единым пакетом, так и отдельно.</p> <p>Базовый пакет STATISTICA Base предоставляет обширные возможности выбора основных типов статистического анализа. Для эффективной работы базового пакета требуется как минимум 256 МВ оперативной памяти. Минимальные требования к скорости процессора – 500 MHz.</p> <p>Модуль <u>Линейные и Нелинейные Модели (Advanced Linear/NonLinearModels)</u> содержит большой набор инструментов для моделирования и прогнозирования, включая возможности автоматического выбора модели и расширенные интерактивные средства визуализации.</p> <p>Модуль <u>Многомерные разведочные технологии анализа (Multivariate Exploratory Techniques)</u> служит для применения разведочного анализа различных типов данных в сочетании с интерактивными средствами визуализации.</p> <p>Помимо общих статистических и графических средств в системе имеются специализированные модули, например, для проведения социологических или биомедицинских исследований, решения технических и промышленных задач, – карты контроля качества, модули анализа процессов и планирования эксперимента.</p> <p>Разработчики STATISTICA утверждают, что пакет может применяться для решения задач в таких областях, как: НИОКР, контроль качества, процесс мониторинга в химической, фармацевтической промышленности и в производстве потребительских товаров:</p> <ul style="list-style-type: none"> <li>гарантийный анализ и приложения для удаленного мониторинга в обрабатывающей промышленности;</li> <li>анализ рисков, сегментация потребителей и оценка кредитоспособности заемщиков в банковской сфере, в сфере предоставления финансовых услуг и в страховой деятельности.</li> </ul> <p>Возможны различные варианты установки пакета в зависимости от целей и задач пользователя:</p> <ul style="list-style-type: none"> <li>однопользовательская версия (<u>Single-User</u>);</li> <li>сетевая версия (<u>Concurrent Network</u>) для работы в локальных вычислительных сетях;</li> <li><u>Enterprise</u> – версия для применения в вычислительных системах и крупных организациях</li> <li><u>Web-Based</u> – версия для использования в крупных сетях через веббраузер.</li> </ul> <p>Пакет предоставляет пользователям следующие возможности статистического анализа данных:</p> <ul style="list-style-type: none"> <li>исследование корреляций между переменными;</li> <li>диаграмма рассеяния, матричная диаграмма рассеяния;</li> <li>быстрые основные статистики и блочные статистики (интерактивные средства, позволяющие одним щелчком мыши вычислять основные статистики и строить графики в любой момент в течение сеанса работы);</li> <li>интерактивный калькулятор вероятностных распределений (позволяет интерактивно исследовать структуру распределений, например, зависимость вероятности от параметров);</li> <li>анализ многомерных откликов, <u>многомерное шкалирование</u>;</li> <li>анализ при помощи временных рядов и прогнозирование временных зависимостей, в том числе анализ сезонных колебаний.</li> </ul>
118	<p><b>Обзор статистического пакета StatGraphics</b> <b>Ответ</b></p> <p>Пакет STATGRAPHICS (STATistical GRAPHICS System)11 – универсальный статистический пакет компании Manugistics Inc12. Первая версия пакета была выпущена в середине 80-х годов. На сегодняшний день разработана 5-я версия пакета Statgraphics Plus 5.1.</p> <p>Достоинства STATGRAPHICS:</p> <ul style="list-style-type: none"> <li>сочетание научных методов обработки разнотипных данных с возможностью создания современной высококачественной интерактивной графики;</li> <li>широкие возможности взаимодействия с другими программными продуктами (электронными таблицами, базами данных);</li> <li>высококачественная двумерная и трехмерная графика,</li> <li>интегрированная графика, предполагающая, что все элементы графических представлений результатов анализа могут быть преобразованы. После завершения процедуры статистического анализа данных можно выбрать графические отображения результатов, релевантные используемой процедуре анализа.</li> </ul> <p>В STATGRAPHICS графика из средства презентации результатов анализа превращается в аналитический инструмент: можно идентифицировать точку на графическом отображении и выяснить ее местонахождение в файле данных или вращать и рассматривать с разных сторон трехмерные изображения, осуществлять разгонку точек на диаграммах рассеяния и т. п.131</p> <p>В STATGRAPHICS существует возможность сохранения результатов работы и создания собственных статистических проектов. После завершения анализа пользователь может сохранить последовательность выбранных методов, параметры статистических процедур, виды графических отображений результатов анализа, табличные формы, комментарии и пр. в отдельном файле. Сохраненную схему анализа можно будет автоматически применять к другому множеству данных.</p>
119	<p><b>Обзор статистического пакета Stadia</b> <b>Ответ</b></p>

	<p style="text-align: center;"><b>Российский статистический пакет STADIA</b></p> <p><b>Пакет STADIA</b> разработан ведущими специалистами Московского государственного университета Ломоносова (главный разработчик - А. П. Кулагин) совместно с НПО "Информатика и компьютеры". Версия пакета была создана в конце 70-х гг. для БЭСМ-6. С тех пор пакет постоянно модифицировался своими функциональными и сервисными возможностями.</p> <p><b>Пакет STADIA</b> является единственным российским статистическим пакетом, представленном на рынке можно отнести к классу универсальных пакетов, то есть в нем представлены все самые распространенные статистического анализа данных от описательной статистики и проверки различных гипотез для временных рядов и контроля качества, а также многомерных (факторный, кластерный, дискриминантный, шкалирование) и непараметрических методов анализа. Таким образом, пакет подходит для решения практически всех задач, встречающихся в исследовании качества жизни.</p> <p><b>Пакет STADIA</b>, в отличие от SAS и SPSS, не поддерживает обработку миллионов наблюдений, но справляется с данными выборочных обследований нескольких сотен или тысяч респондентов. Ориентирован на конкретные статистические расчеты и построение сопутствующих графиков во всей прикладной статистике, снабжая пользователя попутно всей необходимой информацией о работе статистических процедур.</p> <p>В настоящее время пакет используется в учебном процессе и научно-практической работе более 100 университетов России, включая 17 университетов медицинского профиля. Среди пользователей пакета ведущие медицинские центры страны (НИИ им. Сербского, НИИ педиатрии РАМН, НИИ дефектологии, НИИ медико-биологических проблем, НИИ медицинского приборостроения и др.), но и поликлиники, медицинские центры городов: Москвы, Самары, Перми, Тулы, Уфы, Липецка, Архангельска, Кисловодска, Бердянска и др.</p> <p><b>Пакет STADIA</b> простой в освоении, недорогой (профессиональная версия стоит 500 у.е.) и очень удобный инструмент статистического анализа данных ограниченных объемов. Он учитывает уровень подготовки пользователя, позволяет быстро найти необходимый метод обработки, представить результаты анализа в табличной и графической формах и продолжить их оформление средствами среды Windows (текстовых и графических редакторах).</p>
120	<p>Обзор статистического пакета R</p> <p><b>Ответ</b></p> <p>R – статистическая система анализа, созданная Россом Ихаккой и Робертом Гентлеманом (1996, <i>J.Comput. Граф. Stat.</i>, 5: 299-314). R является языком и программным обеспечением; его наиболее замечательные особенности:</p> <ul style="list-style-type: none"> <li>• эффективная обработка данных и простые средства для сохранения результатов,</li> <li>• набор операторов для обработки массивов, матриц, и других сложных конструкций,</li> <li>• большая, последовательная, интегрированная коллекция инструментальных средств для проведения статистического анализа,</li> <li>• многочисленные графические средства,</li> <li>• простой и эффективный язык программирования, который включает много возможностей.</li> </ul> <p>Язык <b>R</b> - рассматривают как диалект языка <b>S</b> созданный AT&amp;T Бэлл Лаборатории. <b>S</b> доступен как программное обеспечение S-PLUS коммерческой системы MathSoft (см.<a href="http://www.splus.mathsoft.com">http://www.splus.mathsoft.com</a> для получения дополнительной информации). Есть существенные различия в концепции <b>R</b> и <b>S</b> (те, кто хочет знать больше об этом может читать статью, написанную Gentleman и Ihaka (1996) или R-FAQ (часто задаваемые вопросы) (<a href="http://cran.r-project.org/doc/FAQ/R-FAQ.html">http://cran.r-project.org/doc/FAQ/R-FAQ.html</a>).</p> <p>R доступен в нескольких формах: исходный текст программ, написанный на C (и некоторые подпрограммы в Fortran77) и в откомпилированном виде.</p> <p>R – язык со многими функциями для выполнения статистического анализа и графического отображения результатов, которые визуализируются сразу же в собственном окне и могут быть сохранены в различных форматах (например, jpg, png, bmp, eps, или wmf под Windows, ps, bmp, pictex под Unix).</p> <p>Результаты статистического анализа могут быть отображены на экране. Некоторые промежуточные результаты (P-values, коэффициент регрессии и т.п.) могут быть сохранены в файле и использоваться для последующего анализа.</p> <p>R – язык, позволяющий пользователю использовать операторы циклов, чтобы последовательно анализировать несколько наборов данных. Также возможно объединить в отдельную программу различные статистические функции, для проведения более сложного анализа.</p>

121	<p>Установка и экосистема R</p> <p><b>Ответ</b>  Рекомендуется устанавливать саму программу (R) и одно из ее графических приложений (RStudio)</p> <p>Ссылка для установки - <a href="https://rstudio.com/products/rstudio/download/">https://rstudio.com/products/rstudio/download/</a></p>  <p>Открытое окно в RStudio</p> <p>В левом верхнем окне пишется код</p> <p>В левом нижнем окне отображаются результаты выполнения кода</p> <p>В правом нижнем окне отображаются графики или справка</p> <p>В левом верхнем окне отображается окружение и текущие значения переменных / объектов, с которыми мы работаем</p>
122	<p>Статистические исследования в R</p> <p><b>Ответ</b></p> <p>Широкий диапазон функций доступен в <b>base</b> пакете. Существует также большое количество других пакетов, которые увеличивают потенциальные возможности <b>R</b>. Они располагаются отдельно и должны быть загружены в память. Исчерпывающий список таких пакетов, вместе с их описаниями, можно найти в Интернете по адресу:  URL: <a href="http://cran.rproject.org/src/contrib/PACKAGES.html">http://cran.rproject.org/src/contrib/PACKAGES.html</a>.</p> <p>В пакете <b>base</b> пакете есть основные статистические модели:</p> <ul style="list-style-type: none"> <li><i>lm</i> линейные модели;</li> <li><i>glm</i> обобщенные линейные модели;</li> <li><i>aov</i>, <i>anova</i> дисперсионный анализ;</li> </ul> <p>В пакете <b>stats</b> пакете есть дополнительные статистические модели, в первую очередь <i>glm</i> – обобщенная линейная модель, позволяющая, например, моделировать логистические или логарифмические зависимости. Пакеты <b>nlme</b>, <b>mgcv</b> позволяют строить нелинейные модели.</p>
123	<p>Анализ данных посредством построения графиков в R</p> <p><b>Ответ</b></p>

	<p>В базовой конфигурации <b>R</b> один из пакетов (графических пакетов несколько, но рассмотрим только один) отвечает за графику — пакет <b>graphics</b>. Выделим три группы входящих в этот пакет функций:</p> <ul style="list-style-type: none"> <li>• функции высокого уровня (высокоуровневые);</li> <li>• функции низкого уровня (низкоуровневые);</li> <li>• интерактивные функции;</li> </ul> <p>Функции первого типа (высокоуровневые) приводят к созданию графического окна, в котором строится заданное изображение.</p> <p>Низкоуровневые функции позволяют изменить (дополнить) уже построенное ранее изображение. Сами они не могут создать графическое окно. Как правило, большинство низкоуровневых функций дублируют аргументы высокоуровневых функций.</p> <p>Наконец, интерактивные функции позволяют по построенному изображению (графику) получить некоторую информацию.</p> <p>В следующих разделах этой главы более подробно будут рассмотрены графические функции.</p>
124	<p>Структуры данных языка <b>R</b> <b>Ответ</b></p> <ul style="list-style-type: none"> <li>• <b>Векторы (vector)</b> — одномерные массивы, состоящие из элементов одного типа данных. Можно выделить числовые, логические и символьные вектора.</li> <li>• <b>Матрицы (matrix)</b> — двумерные массивы, как правило числовые.</li> <li>• <b>Многомерные массивы (array)</b> — массивы, чья размерность больше двух.</li> <li>• <b>Факторы (factor)</b> — структура, полезная при работе с категориальными данными, позволяет определить различные категории данных.</li> <li>• <b>Список (list)</b> — это коллекция объектов, доступ к которым можно осуществить по номеру или имени; список похож на вектор, но его элементы могут быть различных типов.</li> <li>• <b>Таблица данных (data.frame)</b> — наиболее общая структура, используемая при работе в <b>R</b>.</li> </ul>
125	<p>Подбор функции распределения по эмпирическим данным <b>Ответ</b></p>
126	<p>Виды функции Пирсона <b>Ответ</b></p>
127	<p>Параметрические и непараметрические критерии. <b>Ответ</b></p>
128	<p>Способы решения задачи о проверке гипотез о распределении <b>Ответ</b></p>
129	<p>Описание базы данных. Описательные статистики <b>Ответ</b></p>
130	<p>Тесты на симметричность, нормальность распределения, соответствие распределений, проверку гипотез <b>Ответ</b></p>

### 3.5 Вопросы к практическим работам

131	Линейная регрессия. Постановка задачи.
132	Линейная регрессия. Метод наименьших квадратов.
133	Нахождение уравнения регрессии.
134	Оценка тесноты связи уравнения регрессии.
135	Условия Гаусса-Маркова. Доверительные интервалы.
136	Проверка значимости уравнения регрессии, преобразования к линейному виду.
137	Расчет параметров множественной линейной регрессии.
138	Проверка достоверности модели множественной линейной регрессии
139	Регрессия от всех переменных.
140	Регрессия по всем подмножествам.
141	Step-wise регрессия.
142	Случайный лес для отбора признаков.
143	Регрессия с регуляризацией.
144	Метод вычисления процентной важности переменных.
145	MARS-метод.
146	Разделение временного ряда на обучающую и тестовую выборку.
147	Модели ARIMA/SARIMA/ARIMAX
148	Получение прогноза по обученной выборке при анализе временных рядов.
149	Расчет метрик качества прогноза MAE, RMSE, MAPE и выбор наилучшей модели.

### 4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Процедуры оценивания в ходе изучения дисциплины знаний, умений и навыков, характеризующих этапы формирования компетенций, регламентируются положениями:

- П ВГУИТ 2.4.03-2017 Положение о курсовых экзаменах и зачетах;
- П ВГУИТ 4.1.02-2017 Положение о рейтинговой оценке текущей успеваемости.

Оценка по дисциплине средне взвешенная – среднеарифметическое из всех оценок в течение периода изучения дисциплины.

**5 Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания для каждого результата обучения по дисциплине/практике**

Результаты обучения по этапам формирования компетенций	Предмет оценки (продукт или процесс)	Показатель оценивания	Критерии оценивания сформированности компетенций	Шкала оценивания	
				Академическая оценка или баллы	Уровень освоения компетенции
<p><b>УК-1</b> Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач (Умеет анализировать и систематизировать разнородные данные, оценивать эффективность процедур анализа проблем и принятия решений в профессиональной деятельности)</p>					
<b>ЗНАТЬ:</b> методы и принципы сбора разнородных данных	Банк тестовых заданий Вопросы к зачету	Знание теоретических основ дисциплины	Студент ответил не на все вопросы, допустил более 5 ошибок, доля правильных ответов при тестировании менее 50 %	не зачтено	Не освоена (недостаточный)
			Студент ответил не полностью ответил на вопросы, допустил не более 3 ошибок, доля правильных ответов при тестировании более 50%	зачтено	Освоена (базовый)
<b>УМЕТЬ:</b> выполнять систематизацию разнородных данных	Кейс-задание	Знание методики выполнения практических заданий и умение получать решение	Студент не выполнил два задания	не удовлетворительно	Не освоена (недостаточный)
			Студент верно выполнил одно задание и выбрал правильную методику решения хотя бы в одном другом, но допустил ошибки в решении	удовлетворительно	Освоена (базовый)
			Студент верно выполнил два задания и выбрал правильную методику решения, но допустил ошибки в решении третьего задания	хорошо	Освоена (повышенный)
			Студент верно решил все задания или допустил незначительную ошибку в одном из них	отлично	Освоена (повышенный)
<b>ИМЕЕТ НАВЫКИ:</b> анализа данных и оценки эффективности процедур принятия решений	Расчетно-практическая работа	Владение вероятностно-статистическим подходом при решении прикладных задач	Студент выбрал неверную методику решения или не ответил на вопросы или допустил более 5 ошибки в ответе	не удовлетворительно	Не освоена (недостаточный)
			Студент выбрал верную методику решения, представил отчет, имеются значительные замечания по тексту и оформлению работы, ответил на все вопросы, допустил не более 5 ошибки в ответе	удовлетворительно	Освоена (базовый)
			Студент выбрал верную методику	хорошо	Освоена

			решения, представил отчет, имеются незначительные замечания по тексту и оформлению работы, ответил на все вопросы, допустил не более 3 ошибки в ответе		(повышенный)
			Студент выбрал верную методику решения, представил отчет, ответил на все вопросы, допустил не более 1 ошибки в ответе	отлично	Освоена (повышенный)
<p><b>УК-2</b> Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений  (Умеет анализировать альтернативные варианты решений для достижения намеченных результатов; разрабатывать план, определять целевые этапы и основные направления работ)  (Владеет методиками разработки цели и задач проекта; методами оценки продолжительности и стоимости проекта, а также потребности в ресурсах)</p>					
<b>ЗНАТЬ:</b> методики составления альтернативны вариантов и оценки потребности в ресурсах	Банк тестовых заданий Вопросы к зачету	Знание теоретических основ дисциплины	Студент ответил не на все вопросы, допустил более 5 ошибок, доля правильных ответов при тестировании менее 50 %	не удовлетворительно	Не освоена (недостаточный)
			Студент ответил не полностью ответил на вопросы, допустил не более 5 ошибок, доля правильных ответов при тестировании более 50%, но менее 70 %	удовлетворительно	Освоена (базовый)
			Студент ответил на все вопросы допустил более 1, но менее 3 ошибок, доля правильных ответов при тестировании более 70%, но менее 90 %	хорошо	Освоена (повышенный)
			Студент ответил на все вопросы, допустил не более 1 ошибки в ответе, доля правильных ответов при тестировании более 90 %	отлично	Освоена (повышенный)
<b>УМЕТЬ:</b> составлять альтернативные варианты решения задач и применять соответствующие методы оценки потребности в ресурсах	Кейс-задание	Умение анализировать данные и выбирать модели и методы исследования;	Студент выбрал неверную методику решения, не ответил на вопросы или допустил более 5 ошибки в ответе	не удовлетворительно	Не освоена (недостаточный)
			Студент выбрал верную методику решения, оформил отчет, но допустил незначительные ошибки или допустил более 3 ошибок в ответе	удовлетворительно	Освоена (базовый)
			Студент верно выполнил задание, правильно оформил отчет, допустил более 2 ошибок в ответе	хорошо	Освоена (повышенный)
			Студент верно выполнил задание,	отлично	Освоена



			правильно оформил отчет, допустил не более 1 ошибки в ответе		(повышенный)
<b>ИМЕЕТ НАВЫКИ:</b> анализа данных и оценки эффективности процедур принятия решений и оценивания проектов с точки зрения потребности в ресурсах	Расчетно-практическая работа	Исследование и идентификация обобщенной модели множественной регрессии	Студент выбрал неверную методику решения, не ответил на вопросы или допустил более 5 ошибок в ответе	не удовлетворительно	Не освоена (недостаточный)
			Студент выбрал верную методику решения, оформил отчет, но допустил незначительные ошибки или допустил более 3 ошибок в ответе	удовлетворительно	Освоена (базовый)
			Студент верно выполнил задание, правильно оформил отчет, допустил более 2 ошибок в ответе	хорошо	Освоена (повышенный)
			Студент верно выполнил задание, правильно оформил отчет, допустил не более 1 ошибки в ответе	отлично	Освоена (повышенный)
<b>ПКв-4</b> Способность составлять технико-экономическое обоснование проектных решений и техническое задание на разработку информационной системы (Уметь составлять технико-экономического обоснования проектных решений и техническое задание на разработку ИС)					
<b>ЗНАТЬ:</b> приемы и методы составления технико-экономического обоснования	Банк тестовых заданий Вопросы к зачету	Знание теоретических основ дисциплины	Студент ответил не на все вопросы, допустил более 5 ошибок, доля правильных ответов при тестировании менее 50 %	не удовлетворительно	Не освоена (недостаточный)
			Студент ответил не полностью ответил на вопросы, допустил не более 5 ошибок, доля правильных ответов при тестировании более 50%, но менее 70 %	удовлетворительно	Освоена (базовый)
			Студент ответил на все вопросы допустил более 1, но менее 3 ошибок, доля правильных ответов при тестировании более 70%, но менее 90 %	хорошо	Освоена (повышенный)
			Студент ответил на все вопросы, допустил не более 1 ошибки в ответе, доля правильных ответов при тестировании более 90 %	отлично	Освоена (повышенный)
<b>УМЕТЬ:</b> составлять планы проведения технико-экономического обоснования	Кейс-задания	Умение анализировать данные и выбирать модели и методы исследования;	Студент выбрал неверную методику решения, не ответил на вопросы или допустил более 5 ошибок в ответе	не удовлетворительно	Не освоена (недостаточный)
			Студент выбрал верную методику решения, оформил отчет, но допустил незначительные ошибки или допустил более 3 ошибок в ответе	удовлетворительно	Освоена (базовый)
			Студент верно выполнил задание,	хорошо	Освоена (повышен-

			правильно оформил отчет, допустил более 2 ошибок в ответе		ный)
			Студент верно выполнил задание, правильно оформил отчет, допустил не более 1 ошибки в ответе	отлично	Освоена (повышенный)
<b>ИМЕЕТ НАВЫКИ:</b> выполнения технико-экономического обоснования проектных решений	Расчетно-практическая работа	Исследование и идентификация обобщенной модели множественной регрессии	Студент выбрал неверную методику решения, не ответил на вопросы или допустил более 5 ошибки в ответе	не удовлетворительно	Не освоена (недостаточный)
			Студент выбрал верную методику решения, оформил отчет, но допустил незначительные ошибки или допустил более 3 ошибок в ответе	удовлетворительно	Освоена (базовый)
			Студент верно выполнил задание, правильно оформил отчет, допустил более 2 ошибок в ответе	хорошо	Освоена (повышенный)
			Студент верно выполнил задание, правильно оформил отчет, допустил не более 1 ошибки в ответе	отлично	Освоена (повышенный)
<b>ПКв-7</b> Способность настраивать, эксплуатировать и сопровождать информационные системы и сервисы (Участвует в экспертном тестировании ИС на этапе опытной эксплуатации)					
<b>ЗНАТЬ:</b> приемы и метода проведения тестирования систем	Банк тестовых заданий Вопросы к зачету	Знание теоретических основ дисциплины	Студент ответил не на все вопросы, допустил более 5 ошибок, доля правильных ответов при тестировании менее 50 %	не удовлетворительно	Не освоена (недостаточный)
			Студент ответил не полностью ответил на вопросы, допустил не более 5 ошибок, доля правильных ответов при тестировании более 50%, но менее 70 %	удовлетворительно	Освоена (базовый)
			Студент ответил на все вопросы допустил более 1, но менее 3 ошибок, доля правильных ответов при тестировании более 70%, но менее 90 %	хорошо	Освоена (повышенный)
			Студент ответил на все вопросы, допустил не более 1 ошибки в ответе, доля правильных ответов при тестировании более 90 %	отлично	Освоена (повышенный)
<b>УМЕТЬ:</b> применять соответствующие приемы и методы тестирования систем	Кейс-задания	Умение анализировать данные и выбирать модели и методы исследования;	Студент выбрал неверную методику решения, не ответил на вопросы или допустил более 5 ошибки в ответе	не удовлетворительно	Не освоена (недостаточный)
			Студент выбрал верную методику решения, оформил отчет, но допустил незначительные ошибки или допустил более 3 ошибок в ответе	удовлетворительно	Освоена (базовый)

			Студент верно выполнил задание, правильно оформил отчет, допустил более 2 ошибок в ответе	хорошо	Освоена (повышенный)
			Студент верно выполнил задание, правильно оформил отчет, допустил не более 1 ошибки в ответе	отлично	Освоена (повышенный)
<b>ИМЕЕТ НАВЫКИ:</b> проведения тестирования информационных систем	Расчетно-практическая работа	Исследование и идентификация обобщенной модели множественной регрессии	Студент выбрал неверную методику решения, не ответил на вопросы или допустил более 5 ошибки в ответе	не удовлетворительно	Не освоена (недостаточный)
			Студент выбрал верную методику решения, оформил отчет, но допустил незначительные ошибки или допустил более 3 ошибок в ответе	удовлетворительно	Освоена (базовый)
			Студент верно выполнил задание, правильно оформил отчет, допустил более 2 ошибок в ответе	хорошо	Освоена (повышенный)
			Студент верно выполнил задание, правильно оформил отчет, допустил не более 1 ошибки в ответе	отлично	Освоена (повышенный)