

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
Факультет математики и компьютерных наук

УТВЕРЖДАЮ:

Проректор по учебной работе,
качеству образования – первый
проректор

Загуров Т.А.

« 29 » Мая 2020 г.



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.В.13 СТАТИСТИЧЕСКИЕ ПАКЕТЫ

Направление подготовки	02.03.01 Математика и компьютерные науки
Направленность (профиль)	Вычислительные, программные, информационные системы и компьютерные технологии; Алгебра, теория чисел и дискретный анализ; Математическое и компьютерное моделирование
Форма обучения	очная
Квалификация (степень) выпускника	бакалавр

Краснодар 2020

Рабочая программа дисциплины Б1.В.13 Статистические пакеты в соответствии с федеральным государственным образовательным стандартом высшего образования (ФГОС ВО) по направлению подготовки 02.03.01 Математика и компьютерные науки

Программу составила:

О.В. Иванисова, канд. физ.-мат. наук, доцент

Рабочая программа дисциплины утверждена на заседании кафедры вычислительной математики и информатики

протокол № 10 « 15 » апреля 2020 г.

Заведующий кафедрой (разработчик) Гайденко С.В.

Рабочая программа дисциплины утверждена на заседании кафедры вычислительной математики и информатики

протокол № 10 « 15 » апреля 2020 г.

Заведующий кафедрой (выпускающей) Гайденко С.В.

Утверждена на заседании учебно-методической комиссии факультета математики и компьютерных наук

протокол № 2 « 30 » апреля 2020 г.

Председатель УМК факультета Шмалько С.П.

Рецензенты:

Заведующий кафедрой прикладной математики Кубанского государственного университета доктор физико-математических наук профессор Уртенев М.Х.

Доктор экономических наук, кандидат технических наук, профессор кафедры компьютерных технологий и систем КубГАУ Луценко Е.В.

1 Цели и задачи изучения дисциплины (модуля)

11 Цель освоения дисциплины: сформировать у студентов мотивацию к использованию компьютерных статистических пакетов для обработки и анализа статистической информации, подготовить студентов к самостоятельному овладению знаниями, необходимыми для дальнейшей работы в области статистики и компьютерных наук.

12 Задачи дисциплины: освоить основные возможности статистических пакетов для решения статистических задач. Сформировать практические навыки применения статистических пакетов для обработки и анализа статистической информации.

13 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина «Статистические пакеты» относится к части, формируемой участниками образовательных отношений Блока 1 "Дисциплины (модули)" учебного плана.

14 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

Изучение данной учебной дисциплины направлено на формирование у обучающихся общепрофессиональных/профессиональных компетенций (ПК–3, ПК–6)

№ п.п.	Код и наименование компетенции	Индикаторы достижения компетенции		
		знает	умеет	владеет
1.	ПК–3. Способен математически корректно ставить естественнонаучные задачи, знание постановок классических задач математики	Основные понятия теории вероятностей и математической статистики, определения и свойства математических объектов, используемых в этих областях, формулировки утверждений, методы их доказательства, возможные сферы их приложений, основы построения компьютерных моделей	решать задачи теоретического и прикладного характера из различных разделов теории вероятностей и математической статистики, доказывать утверждения, строить модели объектов и понятий	математическим аппаратом теории вероятностей и математической статистики, методами доказательства утверждений в этих областях, навыками алгоритмизации основных задач
2.	ПК–3. Способен использовать методы математического и алгоритмического моделирования при решении теоретических и прикладных задач	Основные методы обработки данных и анализа статистической информации, включенные в статистические пакеты	Использовать статистические пакеты для решения соответствующих задач науки, техники и экономики.	Методами обработки данных и анализа статистической информации с помощью статистических пакетов.

2. Структура и содержание дисциплины

2.1 Распределение трудоёмкости дисциплины по видам работ

Общая трудоёмкость дисциплины составляет 2 зач. ед. (72 часа), их распределение по видам работ представлено в таблице
(для студентов ОФО)

Вид учебной работы		Всего часов	Семестры (часы)		
			7		
Контактная работа, в том числе:		44,2	44,2		
Аудиторные занятия (всего):		40	40		
Занятия лекционного типа		14	14		
Лабораторные занятия		26	26		
Занятия семинарского типа (семинары, практические занятия)					
Иная контактная работа:		4,2	4,2		
Контроль самостоятельной работы (КСР)		4	4		
Промежуточная аттестация (ИКР)		0,2	0,2		
Самостоятельная работа, в том числе:		27,8	27,8		
Проработка учебного (теоретического) материала		8	8		
Выполнение индивидуальных заданий (подготовка сообщений, презентаций)		12	12		
Подготовка к текущему контролю		7,8	7,8		
Контроль:					
Подготовка к экзамену		-	-		
Общая трудоемкость	час.	72	72		
	в том числе контактная работа	44,2	44,2		
	зач. ед	2	2		

2.2 Структура дисциплины

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.

Разделы (темы) дисциплины, изучаемые в 7 семестре (очная форма)

№ раздела	Наименование разделов	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа СРС
			Л	ПЗ	ЛР	
1	2	3	4	5	6	7
1.	Основные понятия теории вероятностей и математической статистики Введение в многомерный статистический анализ.	3	2	0	0	1
2.	Дисперсионный анализ.	3	2	0	0	1
3.	Корреляционный анализ.	2	1	0	0	1
4.	Регрессионный анализ.	4	3	0	0	1
5.	Дискриминантный анализ.	3	2	0	0	1
6.	Кластерный анализ. Факторный анализ	3	2	0	0	1
7.	Приложения многомерного статистического анализа на практике	3	2	0	0	1

№ раздела	Наименование разделов	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	СРС
1	2	3	4	5	6	7
8.	Интегрированная система обработки данных STATISTICA.	3	0	0	2	1
9.	Первичный анализ данных в системе STATISTICA.	3	0	0	2	1
10.	Корреляционный и регрессионный анализы в системе STATISTICA.	5	0	0	3	2
11.	Дисперсионный анализ в системе STATISTICA.	4	0	0	3	1
12.	Дискриминантный анализ в системе STATISTICA.	3	0	0	2	1
13.	Кластерный анализ в системе STATISTICA.	3	0	0	2	1
14.	Факторный анализ в системе STATISTICA.	3	0	0	2	1
15.	Однофакторный дисперсионный анализ в MS Excel	4	0	0	2	2
16.	Двухфакторный дисперсионный анализ без повторений и с повторениями в MS Excel	4	0	0	2	2
17.	Корреляционный анализ в MS Excel	7	0	0	3	4
18.	Регрессионный анализ в MS Excel	7,8	0	0	3	4,8
ИТОГО по разделам дисциплины		67,8	14	0	26	27,8
	Контроль самостоятельной работы (КСР)	4				
	Промежуточная аттестация (ИКР)	0,2				
	Подготовка к текущему контролю					
	Общая трудоемкость по дисциплине	4,2				

Примечание: Л – лекции, ПЗ – практические занятия / семинары, ЛР – лабораторные занятия, СРС – самостоятельная работа студента

2.3 Содержание разделов (тем) дисциплины

2.3.1 Занятия лекционного типа

№	Наименование раздела (темы)	Содержание раздела (темы)	Форма текущего контроля
1	2	3	4
1.	Основные понятия теории вероятностей и математической	Определение понятий: среднее значение, медиана, мода, среднегеометрическая величина, средняя гармоническая,	Устный опрос

	статистики. Введение в многомерный статистический анализ.	дисперсия, среднее квадратичное отклонение, коэффициент вариации, доверительный интервал, асимметрия, эксцесс, квартили, процентиль, корреляция	
2.	Дисперсионный анализ.	Понятие, назначение дисперсионного анализа. Виды дисперсионного анализа. Однофакторный дисперсионный анализ для несвязанных выборок. Однофакторный дисперсионный анализ для связанных выборок. Многофакторный дисперсионный анализ	Устный опрос
3.	Корреляционный анализ.	Цели и задачи корреляционного анализа. Параметрические и непараметрические показатели корреляции	Устный опрос
4.	Регрессионный анализ.	Определение области применения регрессионного анализа. Линейные и нелинейные регрессии. Проверка уравнения регрессии на адекватность	Устный опрос
5.	Дискриминантный анализ.	Понятие, назначение дискриминантного анализа. Модели дискриминантного анализа.	Устный опрос
6.	Кластерный анализ. Факторный анализ	Общая характеристика процедур кластеризации. Анализ качества классификации. Методы кластерного анализа	Устный опрос
7.	Приложения многомерного статистического анализа на практике	Проведение регрессионного, дискриминантного, кластерного анализов на примерах.	Устный опрос

2.3.2 Занятия семинарского типа

Занятия семинарского типа не предусмотрены учебным планом.

2.3.3 Лабораторные занятия

№	Наименование лабораторных работ	Форма текущего контроля
1	3	4
1.	Интегрированная система обработки данных STATISTICA.	ЛР
2.	Первичный анализ данных в системе STATISTICA.	ЛР
3.	Корреляционный и регрессионный анализы в системе STATISTICA.	ЛР
4.	Дисперсионный анализ в системе STATISTICA.	ЛР
5.	Дискриминантный анализ в системе STATISTICA.	ЛР
6.	Кластерный анализ в системе STATISTICA.	ЛР
7.	Факторный анализ в системе STATISTICA.	ЛР
8.	Однофакторный дисперсионный анализ в MS Excel	ЛР
9.	Двухфакторный дисперсионный анализ без повторений и с повторениями в MS Excel	ЛР
10.	Корреляционный анализ в MS Excel	ЛР

11. Регрессионный анализ в MS Excel	ЛР
-------------------------------------	----

Защита лабораторной работы (ЛР), выполнение курсового проекта (КП), курсовой работы (КР), расчетно-графического задания (РГЗ), написание реферата (Р), эссе (Э), коллоквиум (К), тестирование (Т) и т.д.

2.3.4 Примерная тематика курсовых работ (проектов)

Курсовые работы не предусмотрены

2.4 Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

№	Вид СРС	Перечень учебно-методического обеспечения дисциплины по выполнению самостоятельной работы
1	2	3
1	Поиск и анализ литературы и электронных источников информации по заданной проблеме, изучение теоретического материала к лабораторным занятиям, подготовка к зачету	Методические рекомендации по организации самостоятельной работы студентов, утвержденные кафедрой вычислительной математики и информатики, протокол № 14 от 14.06.2017 г.

Учебно-методические материалы для самостоятельной работы обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ) предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа,
- в форме аудиофайла,
- в печатной форме на языке Брайля.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа,
- в форме аудиофайла.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

3. Образовательные технологии

Лабораторные занятия, самостоятельные работы, зачет. В ходе обучения предполагается консультирование студентов с использованием средств информационно-коммуникационных технологий. На лабораторных занятиях выполняются индивидуальные работы. Зачет выставляется после выполнения всех индивидуальных работ.

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

4. Оценочные и методические материалы

4.1 Оценочные средства для текущего контроля успеваемости и промежуточной аттестации

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «название дисциплины».

Оценочные средства включает контрольные материалы для проведения **текущего контроля** в форме практических заданий и **промежуточной аттестации** в форме вопросов и заданий зачету.

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

– при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на экзамене;

– при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;

– при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине (модулю) предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

Структура оценочных средств для текущей и промежуточной аттестации

№ п/п	Контролируемые разделы (темы) дисциплины*	Код контролируемой компетенции (или ее части)	Наименование оценочного средства	
			Текущий контроль	Промежуточная аттестация
1	Основные понятия теории вероятностей и математической статистики. Введение в многомерный статистический анализ.	ПК–3, ПК–6	Вопросы для устного опроса	Вопрос на зачет №1
2	Дисперсионный анализ.	ПК–3, ПК–6	Вопросы для устного опроса	Вопрос на зачет №2
3	Корреляционный анализ.	ПК–3, ПК–6	Вопросы для устного опроса	Вопрос на зачет №3
4	Регрессионный анализ.	ПК–3, ПК–6	Вопросы для устного опроса	Вопрос на зачет №4

5	Дискриминантный анализ.	ПК–3, ПК–6	Вопросы для устного опроса	Вопрос на зачет №5
6	Кластерный анализ. Факторный анализ	ПК–3, ПК–6	Вопросы для устного опроса	Вопрос на зачет №6
7	Приложения многомерного статистического анализа на практике	ПК–3, ПК–6	Вопросы для устного опроса	Вопрос на зачет №7
8	Интегрированная система обработки данных STATISTICA.	ПК–3	Лабораторная работа	Вопрос на зачет №8
9	Первичный анализ данных в системе STATISTICA.	ПК–3, ПК–6	Лабораторная работа	Комплексное задание
10	Корреляционный и регрессионный анализы в системе STATISTICA.	ПК–3, ПК–6	Лабораторная работа	Комплексное задание
11	Дисперсионный анализ в системе STATISTICA.	ПК–3, ПК–6	Лабораторная работа	Комплексное задание
12	Дискриминантный анализ в системе STATISTICA.	ПК–3, ПК–6	Лабораторная работа	Комплексное задание
13	Кластерный анализ системе STATISTICA.	ПК–3, ПК–6	Лабораторная работа	Комплексное задание
14	Факторный анализ системе STATISTICA.	ПК–3, ПК–6	Лабораторная работа	Комплексное задание
15	Однофакторный дисперсионный анализ в MS Excel	ПК–3, ПК–6	Лабораторная работа	Комплексное задание
16	Двухфакторный дисперсионный анализ без повторений и с повторениями в MS Excel	ПК–3, ПК–6	Лабораторная работа	Комплексное задание
17	Корреляционный анализ в MS Excel	ПК–3, ПК–6	Лабораторная работа	Комплексное задание
18	Регрессионный анализ в MS Excel	ПК–3, ПК–6	Лабораторная работа	Комплексное задание

Показатели, критерии и шкала оценки сформированных компетенций

Код и наименование компетенций	Соответствие уровней освоения компетенции планируемым результатам обучения и критериям их оценивания		
	пороговый	базовый	продвинутый
	Оценка		
	Удовлетворительно /зачтено	Хорошо/зачтено	Отлично /зачтено
ПК–3. Способен математически корректно ставить естественнонаучные задачи, знание постановок классических задач математики	<i>Знает</i> - основные понятия теории вероятностей и математической статистики, определения и свойства математических объектов, используемых в этих областях	<i>Знает</i> - основные понятия теории вероятностей и математической статистики, определения и свойства математических объектов, используемых в этих областях, формулировки утверждений, методы их доказательства, возможные сферы их приложений	<i>Знает</i> - основные понятия теории вероятностей и математической статистики, определения и свойства математических объектов, используемых в этих областях, формулировки утверждений, методы их доказательства, возможные сферы их приложений,

			основы построения компьютерных моделей
	<i>Умеет</i> - решать задачи теоретического и прикладного характера из различных разделов теории вероятностей и математической статистики	<i>Умеет</i> - решать задачи теоретического и прикладного характера из различных разделов теории вероятностей и математической статистики, доказывать утверждения	<i>Умеет</i> - решать задачи теоретического и прикладного характера из различных разделов теории вероятностей и математической статистики, доказывать утверждения, строить модели объектов и понятий
	<i>Владеет</i> - математическим аппаратом теории вероятностей и математической статистики	<i>Владеет</i> - математическим аппаратом теории вероятностей и математической статистики, методами доказательства утверждений в этих областях	<i>Владеет</i> - математическим аппаратом теории вероятностей и математической статистики, методами доказательства утверждений в этих областях, навыками алгоритмизации основных задач
ПК–3. Способен использовать методы математического и алгоритмического моделирования при решении теоретических и прикладных задач	<i>Знает</i> - основные методы обработки данных и анализа статистической информации, включенные в статистический пакет STATISTICA	<i>Знает</i> - основные методы обработки данных и анализа статистической информации, включенные в статистический пакет STATISTICA и частично MS Excel	<i>Знает</i> - основные методы обработки данных и анализа статистической информации, включенные в статистический пакет STATISTICA и MS Excel
	<i>Умеет</i> - использовать статистический пакет STATISTICA для решения соответствующих задач науки, техники и экономики	<i>Умеет</i> - использовать статистический пакет STATISTICA и частично MS Excel для решения соответствующих задач науки, техники и экономики	<i>Умеет</i> - использовать статистический пакет STATISTICA и MS Excel для решения соответствующих задач науки, техники и экономики
	<i>Владеет</i> - методами обработки данных и анализа статистической информации с помощью статистического пакета STATISTICA	<i>Владеет</i> - методами обработки данных и анализа статистической информации с помощью статистического пакета STATISTICA и частично MS Excel	<i>Владеет</i> - методами обработки данных и анализа статистической информации с помощью статистического пакета STATISTICA и MS Excel

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

Вопросы для устного опроса

1. STATISTICA. Корреляционный анализ.
 - 1.1. Вывести Подробную таблицу результатов. Какая корреляционная связь Значимая?
 - 1.2. Вывести Матрицу диаграмм рассеивания для всех переменных. Что показывает диаграмма рассеивания для частного случая.
2. STATISTICA. Регрессионный анализ.
 - 2.1. Объяснить получившиеся значения коэффициента детерминации и значения F-Фишера.
 - 2.2. Записать уравнение регрессии в явном виде, с пояснением значимости коэффициентов уравнения регрессии.
 - 2.3. Построить график рассеивания исходных данных. Исключить из графика с помощью средства Кисть данные по 3 строке.

- 2.4. Построить график Наблюдаемые переменные – остатки. Объяснить смысл данного графика.
- 2.5. Построить график Наблюдаемые и предсказанные переменные – остатки. Объяснить смысл данного графика.
- 2.6. Определить значение y при заданном значении x .
3. STATISTICA. Дисперсионный анализ.
 - 3.1. Как проводить дисперсионный анализ в программе STATISTICA
 - 3.2. Как просмотреть анализ результатов – просмотр всех эффектов
 - 3.3. Как просмотреть анализ результатов – просмотр заданных эффектов
 - 3.4. Как просмотреть анализ результатов – оценка качества модели
 - 3.5. Как просмотреть анализ результатов – анализ контрастов
4. STATISTICA. Дискриминантный анализ.
 - 4.1. Записать правило отнесения потребителя к одному из выделенных классов.
 - 4.2. Определить к какому классу относится потребитель с данными: 25; 4,3; 80; 9; 2; 1.
 - 4.3. Построить диаграмму рассеивания канонических значений.
 - 4.4. Объяснить получившиеся значения Ламбда Уилкса и значения F-Фишера.
 - 4.5. Провести дискриминантный анализ для строк 11-20.
5. STATISTICA. Кластерный анализ.
 - 5.1. Объяснить построенную дендрограмму, матрицу расстояний.
 - 5.2. Провести кластерный анализ методом «K-средних» с параметрами: 3 кластера, Наблюдения строки.
 - 5.3. Показать элементы кластеров. Объяснить график средних.
6. STATISTICA. Факторный анализ.
 - 6.1. Отобразить таблицу результатов с корреляциями. Что можно сказать о получившихся значениях?
 - 6.2. Отобразить таблицу Собственные значения. Что можно сказать о получившихся значениях?
7. Однофакторный дисперсионный анализ в MS Excel.
 - 7.1. Какие функции используются для построения данного анализа?
 - 7.2. Описать алгоритм построения модели в MS Excel.
8. Двухфакторный дисперсионный анализ без повторений и с повторениями в MS Excel.
 - 8.1. Какие функции используются для построения данного анализа?
 - 8.2. Описать алгоритм построения модели в MS Excel.
9. Корреляционный анализ в MS Excel. Какие функции используются для построения данного анализа?
 - 9.1. Какие функции используются для построения данного анализа?
 - 9.2. Описать алгоритм построения модели в MS Excel.
10. Регрессионный анализ в MS Excel. Какие функции используются для построения данного анализа?
 - 10.1. Какие функции используются для построения данного анализа?
 - 10.2. Описать алгоритм построения модели в MS Excel.

Перечень компетенций (части компетенции), проверяемых оценочным средством:
ПК-3, ПК-6

Задачи для лабораторных работ

Боровиков В.П. Популярное введение в современный анализ данных в системе STATISTICA. Учебное пособие для вузов : учеб. пособие — Москва: Горячая линия-Телеком, 2013. — 288 с. <https://e.lanbook.com/book/11828>

Перечень компетенций (части компетенции), проверяемых оценочным средством:
ПК-3, ПК-6

Зачетно-экзаменационные материалы для промежуточной аттестации (экзамен/зачет)

Пример комплексного задания промежуточной аттестации

Задача 1. По семи территориям Уральского района известны значения двух признаков.

Район	Расходы на покупку продовольственных товаров в общих расходах, %, у	Среднедневная заработная плата одного работающего, руб., х
Удмуртская респ.	68,8	45,1
Свердловская обл.	61,2	59,0
Башкортостан	59,9	57,2
Челябинская обл.	56,7	61,8
Пермская обл.	55,0	58,8
Курганская обл.	54,3	47,2
Оренбургская обл.	49,3	55,2

Требуется:

1. Провести линейный регрессионный анализ зависимости у от х в системе STATISTICA.
2. Объяснить получившиеся значения коэффициента детерминации и значения F-Фишера.
3. Записать уравнение регрессии в явном виде, с пояснением значимости коэффициентов уравнения регрессии.
4. Построить график рассеивания исходных данных. Исключить из графика с помощью средства Кисть данные по 3 строке.
5. Построить график Наблюдаемые переменные – остатки. Объяснить смысл данного графика.
6. Построить график Наблюдаемые и предсказанные переменные – остатки. Объяснить смысл данного графика.
7. Определить значение у при значении х 63.
8. По исходным данным постройте информационную модель для линейного регрессионного анализа в MS Excel.

Задача 2. Исследуется рынок потребителей услуги «Подключение к сети Интернет». С целью выбора оптимального сегмента для позиционирования услуги на рынке необходимо разделить потребителей на группы в соответствии с принадлежностью к профессии.

Исходные данные для проведения анализа представлены в таблице ниже:

Номер п/п	возраст	стаж	кол-во час.	доход	профессия	пол
1	34	1,2	16,1	5,6	1	2
2	29	5,0	22,9	8,8	1	2
3	24	5,1	82,3	9,7	2	2
4	41	0,4	6,4	7,6	0	1
5	29	1,5	14,3	10,6	1	1
6	25	5,2	83,6	8,1	2	2
7	32	2,7	23,2	11,2	1	2
8	27	2,7	23,4	8,6	1	2
9	20	1,5	15,1	9,7	1	1
10	34	1,7	15,2	10,2	1	2
11	35	0,5	6,4	13,8	0	2
12	25	4,7	79,6	8,3	2	2
13	41	0,4	6,4	12,6	0	1

14	24	5,4	35,0	9,5	2	2
15	44	0,4	6,5	11,7	0	1
16	27	5,2	81,4	12,1	2	2
17	39	0,6	6,3	8,3	0	1
18	40	0,6	5,9	8,8	0	2
19	29	2,2	24,3	11,9	1	1
20	29	2,5	24,0	5,3	1	2

Требуется:

1. Провести дискриминантный анализ и записать правило отнесения потребителя к одному из выделенных классов.
2. Определить к какому классу относится потребитель с данными: 25; 4,3; 80; 9; 2; 1.
3. Построить диаграмму рассеивания канонических значений.
4. Объяснить получившиеся значения Ламбда Уилкса и значения F-Фишера.
5. Провести дискриминантный анализ для строк 11-20.
6. По исходным данным постройте информационную модель для дискриминантного анализ в MS Excel.

Задача 3.

	Age	Weight (lbs.)	Body Length (inches)
Tori	3	8,52	9,3
Cheshire	4	9,23	19,5
Fluffy	5	6,97	16,2
Cotton	3	8,5	14
Isabelle	5	13,5	16,25
Salem	5	9,7	17,1
Muffin	2	10,2	17,5
Hobbes	8	8,4	16,6
Sylvester	7	6,75	18,6
Emma Lu	1	22	30,3
Angel	6	5,2	15,51
Chester	3	5,1	12,1
Angel	4	4,1	14,5
Mr. Magoo	3	7,9	13,7

Требуется:

1. Провести кластерный анализ методом «Иерархическая классификация» с параметрами: Наблюдения строки, Евклидово расстояние, метод одиночной связи.
2. Объяснить построенную дендрограмму, матрицу расстояний.
3. Провести кластерный анализ методом «K-средних» с параметрами: 3 кластера, Наблюдения строки.
4. Показать элементы кластеров. Объяснить график средних.

Перечень компетенций (части компетенции), проверяемых оценочным средством:
ПК-3, ПК-6

Вопросы для подготовки к зачету

1. Определение понятий описательной статистики: среднее значение, медиана, мода, среднегеометрическая величина, средняя гармоническая, дисперсия, среднеквадратичное

отклонение, коэффициент вариации, доверительный интервал, асимметрия, эксцесс, квартили, процентиль, корреляция.

2. Понятие, назначение дисперсионного анализа. Виды дисперсионного анализа. Однофакторный дисперсионный анализ для несвязанных выборок. Однофакторный дисперсионный анализ для связанных выборок. Многофакторный дисперсионный анализ

3. Цели и задачи корреляционного анализа. Параметрические и непараметрические показатели корреляции

4. Определение области применения регрессионного анализа. Линейные и нелинейные регрессии. Проверка уравнения регрессии на адекватность

5. Понятие, назначение дискриминантного анализа. Модели дискриминантного анализа.

6. Общая характеристика процедур кластеризации. Анализ качества классификации. Методы кластерного анализа.

7. Приведите примеры использования различных методов статистического анализа, применяемых для исследования статистических данных.

8. Опишите назначение и основной функционал системы STATISTICA.

Перечень компетенций (части компетенции), проверяемых оценочным средством:
ПК-3, ПК-6

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

Методические рекомендации, определяющие процедуры оценивания на зачете:

Согласно учебному плану дисциплины студенты сдают зачет в 7-м семестре.

Зачет выставляется при достаточно полном знании материала учебной программы, отсутствии грубых ошибок при его изложении и ответах на вопросы, умении на основе теоретических знаний делать практические выводы, применять полученные знания при решении задач.

Преподаватель учитывает работу студента в течение семестра: успешно выполненная контрольная работа положительно влияет на оценивание знаний студента.

Зачет не может быть получен при незнании основных вопросов курса, неумении на основе теоретических знаний делать практические выводы и применять полученные знания при решении задач.

5. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)

5.1 Основная литература:

1. Боровиков В.П. Популярное введение в современный анализ данных в системе STATISTICA. Учебное пособие для вузов : учеб. пособие — Москва: Горячая линия-Телеком, 2013. — 288 с. <https://e.lanbook.com/book/11828>

Кобзарь А. И. Прикладная математическая статистика: для инженеров и научных работников: учебное пособие /Кобзарь А. И. - М. : ФИЗМАТЛИТ, 2012. - 816 с. - <https://e.lanbook.com/book/59747>

3. Статистический анализ данных. STATISTICA 6 [Текст] : учебное пособие для студентов вузов / А. А. Халафян. - [3-е изд.]. - М. : [БИНОМ-Пресс], 2007. - 508 с. : ил. - (Учебник). - Библиогр.: с. 507-508. - ISBN 9785951802156

Для освоения дисциплины инвалидами и лицами с ограниченными возможностями здоровья имеются издания в электронном виде в электронно-библиотечных системах «Лань» и «Юрайт».

5.2 Дополнительная литература:

1. Математическая статистика. Примеры и задачи : учебное пособие. - Новосибирск : НГТУ, 2011. - 84 с. <http://biblioclub.ru/index.php?page=book&id=229133>

2. Балдин, К. В. Теория вероятностей и математическая статистика : учебник / К. В. Балдин, В. Н. Башлыков, А. В. Рукосуев. - 2-е изд. - Москва : Дашков и К°, 2016. - 472 с. - <http://biblioclub.ru/index.php?page=book&id=453249>

5.3. Периодические издания:

Периодические издания — не предусмотрены.

6. Методические указания для обучающихся по освоению дисциплины (модуля)

Для сдачи зачета надо изучить теоретический материал таблицы п.2.3.1. Также студент должен научиться выполнять практические задания по темам этих разделов на лабораторных занятиях. Самостоятельная работа студента включает в себя подготовку к лабораторным занятиям и зачету. Эти виды самостоятельной работы студентов контролируется в ходе проверки домашних заданий и зачета. Теоретические вопросы к зачету приведены в пункте 4. Зачет выставляется после успешного выполнения лабораторных работ.

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья большое значение имеет индивидуальная учебная работа (консультации) – дополнительное разъяснение учебного материала.

Индивидуальные консультации по предмету являются важным фактором, способствующим индивидуализации обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или лицом с ограниченными возможностями здоровья.

7. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю)

7.1 Перечень информационно-коммуникационных технологий
образовательные ресурсы интернета, Интернет-коммуникации (электронная почта)

7.2 Перечень лицензионного и свободно распространяемого программного обеспечения
STATISTICA, MS Office Excel

7.3 Перечень современных профессиональных баз данных и информационных справочных систем

1. Электронная библиотечная система eLIBRARY.RU (<https://elibrary.ru>)
2. Электронно-библиотечная система издательства «Лань» (<http://e.lanbook.com>)
3. Электронная библиотечная система «Университетская библиотека ONLINE» (www.biblioclub.ru)

8. Материально-техническое обеспечение по дисциплине (модулю)

№	Вид работ	Наименование учебной аудитории, ее оснащенность оборудованием и техническими средствами обучения
1.	Лекционные занятия	Лекционная аудитория, оснащенная доской и презентационной техникой (проектор, экран, компьютер/ноутбук) и соответствующим программным обеспечением (ПО)

2.	Лабораторные занятия	Специальное помещение, оснащенное компьютерной техникой (системный блок, монитор, клавиатура, мышь) и соответствующим программным обеспечением (ПО)
3.	Групповые (индивидуальные) консультации	Специальное помещение, оснащенное компьютерной техникой и соответствующим программным обеспечением (ПО)
4.	Текущий контроль, промежуточная аттестация	Специальное помещение, оснащенное компьютерной техникой и соответствующим программным обеспечением (ПО)
5.	Самостоятельная работа	Кабинет для самостоятельной работы, оснащенный компьютерной техникой с возможностью подключения к сети «Интернет», программой экранного увеличения и обеспеченный доступом в электронную информационно-образовательную среду университета.