Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Сибирский государственный индустриальный университет»

Кафедра прикладной математики и информатики

УТВЕРЖДАЮ
Директор института
информационных технологий и
автоматизированных систем
_____ А.В. Зимин
подпись
«____» _____ 20___ г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Прикладная статистика

01.03.02 «Прикладная математика и информатика»
(направленность (профиль): «Математическое моделирование и интеллектуальный анализ данных»)

Квалификация выпускника
Бакалавр

Форма обучения
Очная форма

Срок обучения: 4 года

Год начала подготовки 2025

Новокузнецк
2025

СибГИУ
Подлинник электронного документа хранится в системе электронного согласования документов СибГИУ
ДОКУМЕНТ ПОДПИСАН ЭП
Сертификат: daba56d7b892efca3bf5c04ed0251c98b19b099c
Владелец: Директор ИИТиАС Зимин Алексей Валерьевич
Действителен: с 15 января 2025 г. по 10 апреля 2026 г.
Дата и время: 14-04-2025 09:44:31 ПРОВЕРИТЬ ПОДПИСЬ

1 Цели и задачи освоения учебной дисциплины

Целями учебной дисциплины являются:

– освоение основных разделов прикладной статистики и способов применение статистических методов в инженерной практике и научной деятельности;
– формирование понятийного аппарата о принципах, приемах и методах, предназначенных для систематизации и обработки статистических данных с целью их удобного представления, интерпретации и получения практических выводов.

Задачами учебной дисциплины являются:

– Изучение методов статистики, точечных и интервальных оценок, статистических критериев, методов классического регрессионного анализа;
– формирование навыков решения задач анализа и обработки данных, необходимых в научно-практической деятельности прикладного характера.

2 Место учебной дисциплины в структуре ООП по направлению подготовки (специальности)

Учебная дисциплина относится к учебным дисциплинам части, формируемой участниками образовательных отношений **Блока 1 «Дисциплины (модули)»** ООП по направлению подготовки (специальности) 01.03.02 «Прикладная математика и информатика».

Учебная дисциплина базируется на предварительном усвоении обучающимися учебных дисциплин:

– Математика;
– Специальные главы математики;
– Теория вероятностей и математическая статистика.

Учебная дисциплина дополняет знания и умения, получаемые по одновременно изучаемым и последующим учебным дисциплинам:

– Теория прогнозирования;
– Анализ временных рядов и прогнозирование.

3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

– Профессиональные компетенции

Наименование категории (группы) ПК	Код и наименование ПК	Код и наименование индикатора достижения ПК	Планируемые результаты обучения
	ПК-2: Способен проводить	ПК-2.1 Выбирает методы и инструментальные	– знать: методы и инструментальные средства анализа

	аналитические исследования с применением технологии больших данных в соответствии с требованиями заказчика	средства анализа больших данных для проведения аналитических работ	больших данных. – уметь: применять методы и средства прикладной статистики для проведения аналитических работ.
		ПК-2.3 Выбирает средства представления результатов аналитики больших данных	– знать: правила оформления результатов аналитических исследований. – уметь: самостоятельно выбирать методы и средства для аналитики данных.

4 Объем и содержание учебной дисциплины

Учебные занятия по учебной дисциплине проводятся в форме контактной работы и в форме самостоятельной работы обучающихся.

Контактная работа включает в себя занятия лекционного типа (лекции), занятия семинарского типа (семинары, практические занятия, практикумы), промежуточную аттестацию обучающихся и иные формы взаимодействия обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации ООП на иных условиях, в том числе при проведении промежуточной аттестации обучающихся. Контактная работа может проводиться с применением электронного обучения, дистанционных образовательных технологий.

Объем учебной дисциплины

Семестр / курс		***ИТОГО***	***7 семестр***	***8 семестр***
Форма промежуточной аттестации			*экзамен*	*экзамен*
Трудоёмкость	*академ. час.*	***288***	*180*	*108*
	зачетных единиц	***8***	*5*	*3*
Лекции, *академ. час.*		***50***	*32*	*18*
в форме практической подготовки		***0***	*0*	*0*
Лабораторные работы, *академ. час.*		***0***	*0*	*0*
в форме практической подготовки		***0***	*0*	*0*
Практические занятия, *академ. час.*		***50***	*32*	*18*
в форме практической подготовки		***0***	*0*	*0*
Курсовая работа / проект, *академ. час.*		***0***	*0*	*0*
в форме практической подготовки		***0***	*0*	*0*
Консультации, *академ. час.*		***0***	*0*	*0*
в форме практической подготовки		***0***	*0*	*0*
Самостоятельная работа, *академ. час.*		***143***	*89*	*54*

в форме практической подготовки	***0***	*0*	*0*
Контроль, *академ. час.*	***45***	*27*	*18*
в форме практической подготовки	***0***	*0*	*0*

Содержание учебной дисциплины

Раздел 1 Характеристика двумерной случайной величины;

Тема 1.1 Дискретные двумерные случайные величины (Понятия переменной, функциональной и статистической зависимости, случайной величины. Двумерная случайная величина, дискретные величины, функция распределения);

Тема 1.2 Непрерывные двумерные случайные величины (Непрерывные случайные величины, плотности распределения вероятностей случайных величин, функция распределения случайной величины);

Раздел 2 Парный корреляционный анализ;

Тема 2.1 Парный коэффициент корреляции (Оценка параметров двумерной нормально распределенной генеральной совокупности (генеральные средняя, дисперсия, коэффициент корреляции). Плотность двумерного нормального закона распределения случайной величины. Коэффициент корреляции);

Тема 2.2 Проверка значимости истинного коэффициента корреляции (Нулевая гипотеза. Значимость коэффициента корреляции. Распределение Стьюдента);

Тема 2.3 Интервальное оценивание коэффициента корреляции (Доверительный интервал, выборочный коэффициент корреляции, закон распределения. z - преобразование Фишера);

Раздел 3 Множественный корреляционный анализ;

Тема 3.1 Частная корреляция (Корреляционная матрица, частные коэффициенты корреляции и их свойства. Определение значимости частных коэффициентов корреляции. Коэффициент множественной детерминации);

Тема 3.2 Множественная корреляция (Определение тесноты линейной корреляционной связи. Множественный коэффициент корреляции. Матрица парных коэффициентов корреляции. Множественный коэффициент детерминации);

Раздел 4 Парный регрессионный анализ;

Тема 4.1 Уравнение парной регрессии в стандартизованном масштабе (Неслучайная составляющая, параметры, случайный член. Основные причины включения случайного члена. Аппроксимация набора наблюдений посредством минимизации функционала, необходимое условие экстремума);

Тема 4.2 Оценка коэффициентов парной регрессии (Коэффициенты регрессии. Оценка параметров уравнения регрессии: тесноты связи, качества уравнение регрессии, статистической

надежности результатов регрессионного моделирования, прогнозного значения.);

Раздел 5 Множественный регрессионный анализ;

Тема 5.1 Уравнение множественной регрессии (Цели регрессионного анализа. Оценка параметров регрессии. Уравнение регрессии в стандартизованном масштабе. Расчет частных коэффициентов эластичности. Мультиколлинеарность факторов);

Тема 5.2 Оценка значимости уравнения множественной регрессии (Критерии оценок: несмещенность, эффективность, состоятельность. Оценка регрессии на основе минимизации суммы квадратов остатков. Гомоскедастичная и гетероскедастичная дисперсия остатков. t-критерий Стьюдента, F-критерий Фишера);

Тема 5.3 Определение вектора оценок коэффициентов множественной регрессии (Матричный способ определения вектора оценок коэффициентов регрессии. Оценка статистической значимости эмпирических коэффициентов регрессии. Проверка качества построенного уравнения регрессии);

Раздел 6 Многомерный статистический анализ;

Тема 6.1 Многомерные статистические методы (Кросс-табуляция, дисперсионный, ковариационный, дискриминантный, канонический, факторный, кластерный анализ. Многомерное шкалирование);

Тема 6.2 Построение уравнения множественной регрессии по главным компонентам (Уравнение множественной регрессии, уменьшение размерности модели методом главных компонент);

Тема 6.3 Факторный анализ идентификации факторов (Фактор, собственные значения фактора. основные задачи факторного анализа. Кластерный анализ корреляций).

5 Перечень тем лекций

№ раздела / темы дисциплины	Темы лекций	Трудоемкость, *академ.час*	
		всего	**в форме практической подготовки**
Раздел 1; Тема 1.1.	Дискретные двумерные случайные величины	4	
Раздел 1; Тема 1.2.	Непрерывные двумерные случайные величины	4	
Раздел 2; Тема 2.1.	Парный коэффициент корреляции	4	
Раздел 2; Тема 2.2.	Проверка значимости истинного коэффициента корреляции	2	
Раздел 2; Тема 2.3.	Интервальное оценивание коэффициента корреляции	2	
Раздел 3; Тема 3.1.	Частная корреляция	4	

Раздел 3; Тема 3.2.	Множественная корреляция	4	
Раздел 4; Тема 4.1.	Уравнение парной регрессии в стандартизованном масштабе	4	
Раздел 4; Тема 4.2.	Оценка коэффициентов парной регрессии	4	
Раздел 5; Тема 5.1.	Уравнение множественной регрессии	2	
Раздел 5; Тема 5.2.	Оценка значимости уравнения множественной регрессии	2	
Тема 5.1; Тема 5.3.	Определение вектора оценок коэффициентов множественной регрессии	4	
Раздел 6; Тема 6.1.	Многомерные статистические методы	2	
Раздел 6; Тема 6.2.	Построение уравнения множественной регрессии по главным компонентам	4	
Раздел 6; Тема 6.3.	Факторный анализ идентификации факторов	4	
Итого:		**50**	**0**

6 Перечень тем практических занятий (семинаров)

№ раздела / темы дисциплины	**Темы практических занятий (семинаров)**	**Трудоемкость, *академ.час***	
		всего	**в форме практической подготовки**
Раздел 1; Тема 1.1.	Дискретные двумерные случайные величины	4	
Раздел 1; Тема 1.2.	Непрерывные двумерные случайные величины	4	
Раздел 2; Тема 2.1.	Парный коэффициент корреляции	4	
Раздел 2; Тема 2.2.	Проверка значимости истинного коэффициента корреляции	2	
Раздел 2; Тема 2.3.	Интервальное оценивание коэффициента корреляции	2	
Раздел 3; Тема 3.1.	Частная корреляция	4	
Раздел 3; Тема 3.2.	Множественная корреляция	4	
Раздел 4; Тема 4.1.	Уравнение парной регрессии в стандартизованном масштабе	4	
Раздел 4; Тема 4.2.	Оценка коэффициентов парной регрессии	4	

Раздел 5; Тема 5.1.	Уравнение множественной регрессии	2	
Раздел 5; Тема 5.2.	Оценка значимости уравнения множественной регрессии	2	
Раздел 5; Тема 5.3.	Определение вектора оценок коэффициентов множественной регрессии	4	
Раздел 6; Тема 6.1.	Многомерные статистические методы	2	
Раздел 6; Тема 6.2.	Построение уравнения множественной регрессии по главным компонентам	4	
Раздел 6; Тема 6.3.	Факторный анализ идентификации факторов	4	
Итого:		**50**	**0**

7 Перечень тем лабораторных работ

№ раздела / темы дисциплины	**Темы лабораторных работ**	**Трудоемкость, *академ.час***	
		всего	**в форме практической подготовки**
	Отсутствуют		
Итого:		**0**	**0**

8 Перечень тем курсовых работ (проектов)

№ раздела / темы дисциплины	**Темы курсовых работ (проектов)**	**Трудоемкость, *академ.час***	
		всего	**в форме практической подготовки**
	Отсутствуют		
Итого:		**0**	**0**

9 Виды самостоятельной работы

№ раздела / темы дисциплины	**Виды самостоятельной работы**	**Трудоемкость, *академ.час***	
		всего	**в форме практической подготовки**
Раздел 1.	1. Изучение лекционного материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования.	29	
Раздел 2.	1. Изучение лекционного материала; 2. Подготовка к практическому занятию; 3. Прохождение	30	

	тестирования.		
Раздел 3.	1. Изучение лекционного материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования.	30	
Раздел 4.	1. Изучение лекционного материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования.	18	
Раздел 5.	1. Изучение лекционного материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования.	18	
Раздел 6.	1. Изучение лекционного материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования.	18	
Контроль	*Подготовка к экзамену (7 семестр)*	27	
Контроль	*Подготовка к экзамену (8 семестр)*	18	
Итого:		**188**	**0**

10 Матрица компетентностных задач по учебной дисциплине

Контролируемые разделы (темы) дисциплины	Контролируемые компетенции	Оценочные средства
Раздел 1.	ПК-2	1. Вопросы для устного обсуждения (опроса); 2. Наблюдение за работой на практическом занятии; 3. Тестовые задания
Раздел 2.	ПК-2	1. Вопросы для устного обсуждения (опроса); 2. Наблюдение за работой на практическом занятии; 3. Тестовые задания
Раздел 3.	ПК-2	1. Вопросы для устного обсуждения (опроса); 2. Наблюдение за работой на практическом занятии; 3. Тестовые задания
Раздел 4.	ПК-2	1. Вопросы для устного

		обсуждения (опроса); 2. Наблюдение за работой на практическом занятии; 3. Тестовые задания
Раздел 5.	ПК-2	1. Вопросы для устного обсуждения (опроса); 2. Наблюдение за работой на практическом занятии; 3. Тестовые задания
Раздел 6.	ПК-2	1. Вопросы для устного обсуждения (опроса); 2. Наблюдение за работой на практическом занятии; 3. Тестовые задания
Экзамен (7 семестр).	ПК-2	1. Вопросы к экзамену; 2. Задачи к экзамену
Экзамен (8 семестр).	ПК-2	1. Вопросы к экзамену; 2. Задачи к экзамену

11 Входной контроль

11.1 Цель входного контроля − определить начальный уровень подготовленности обучающихся и выстроить индивидуальную траекторию обучения. В условиях личностно-ориентированной образовательной среды результаты входного оценивания обучающегося используются как начальные значения в индивидуальном профиле академической успешности обучающегося.

11.2 Описание оценочных средств

Форма проведения входного контроля – тестирование.

Длительность тестирования – 30 минут.

Тестирование проводится в течение первых двух недель изучения учебной дисциплины в СУО «Moodle».

Для проведения тестирования используются тестовые задания для входного контроля (для ответа на 1 вопрос, как правило, выделяется 1 минута).

Шкала оценивания результатов тестирования

За каждый правильный ответ выставляется 1 балл. Оценка выставляется в зависимости от процента правильных ответов:

Оценка формируется в соответствии с критериями таблицы:

Процент результативности (правильных ответов)	Оценка
от 50 до 100	«зачтено»
менее 50	«не зачтено»

Основываясь на результатах тестирования, формируется индивидуальная траектория обучающегося (группы обучающихся), исходя из следующих соображений:

– 50 – 100 % – обучающийся готов к изучению учебной дисциплины без дополнительного повторения изученного ранее материала;

– 49% и менее – обучающийся имеет слабую подготовку к изучению учебной дисциплины, необходимы корректирующие занятия и консультации.

Тестовые задания для входного контроля знаний обучающихся приведены в приложении.

12 Текущий контроль

12.1 Текущий контроль знаний используется для оперативного и регулярного управления учебной деятельностью (в том числе самостоятельной) обучающихся. Данный вид контроля стимулирует у обучающихся стремление к систематической самостоятельной работе по изучению учебной дисциплины и позволяет оценить степень освоения изученного материала.

Текущий контроль успеваемости осуществляется в течение семестра, в ходе повседневной учебной работы по индивидуальной инициативе педагогического работника, в том числе посредством использования СУО «Moodle».

Формы проведения текущего контроля – вопросы для устного обсуждения (опроса); наблюдение за работой на практическом занятии; тестовые задания.

12.2 Описание оценочных средств

Шкалы оценивания (методики оценки)

12.2.1 Рекомендации по оцениванию устных ответов обучающихся

С целью контроля и подготовки обучающихся к изучению новой темы в начале занятия лекционного типа (лекции) педагогическим работником проводится индивидуальный или фронтальный устный опрос по выполненным заданиям предыдущего раздела (темы).

Критерии оценки:

– правильность ответа по содержанию задания (учитывается количество и характер ошибок при ответе);

– полнота и глубина ответа (учитывается количество усвоенных фактов, понятий и т.п.);

– сознательность ответа (учитывается понимание излагаемого материала);

– логика изложения материала (учитывается умение строить целостный, последовательный рассказ, грамотно пользоваться специальной терминологией);

– рациональность использованных приемов и способов решения поставленной учебной задачи (учитывается умение использовать наиболее прогрессивные и эффективные способы достижения цели);

– своевременность и эффективность использования наглядных пособий и технических средств при ответе (учитывается умение грамотно и с пользой применять наглядность и демонстрационный опыт при устном ответе);

– использование дополнительного материала (обязательное условие);

– рациональность использования времени, отведенного на задание (не одобряется затянутость выполнения задания, устного ответа во времени).

Оценка «отлично» ставится, если обучающийся:

– полно и аргументировано отвечает по содержанию задания;

– обнаруживает понимание материала, может обосновать свои суждения, применить знания на практике, привести необходимые примеры, в том числе самостоятельно составленные;

– излагает материал последовательно и правильно.

Оценка «хорошо» ставится, если обучающийся дает ответ, удовлетворяющий тем же требованиям, что и для оценки «отлично», но допускает 1 – 2 ошибки, которые сам же исправляет.

Оценка «удовлетворительно» ставится, если обучающийся обнаруживает знание и понимание основных положений данного задания, но:

– излагает материал неполно и допускает неточности в определении понятий или формулировке правил;

– не умеет достаточно глубоко и доказательно обосновать свои суждения и привести свои примеры;

– излагает материал непоследовательно и допускает ошибки.

Оценка «неудовлетворительно» ставится, если обучающийся обнаруживает незнание ответа на соответствующее задание, допускает ошибки в формулировке определений и правил, искажающие их смысл, беспорядочно и неуверенно излагает материал. Оценка «неудовлетворительно» отмечает такие недостатки в подготовке обучающегося, которые являются серьезным препятствием к успешному освоению последующего материала.

12.2.2 Рекомендации по оцениванию результатов тестирования обучающихся

Тестирование проводится по изученным разделам (темам) учебной дисциплины для выявления уровня знаний и умений обучающегося, мотивирования обучающегося к активизации самостоятельной работы

по усвоению учебного материала, выявления и устранения пробелов в знаниях, повышения дисциплины и организации учебной деятельности обучающегося.

Тестирование проводится в СУО «Moodle» в рамках часов, отведенных на самостоятельную работу обучающихся. Для тестирования используются вопросы из базы типовых тестовых заданий по изученным разделам (темам) учебной дисциплины. Обучающийся выбирает один ответ из предложенных вариантов ответа. В случае выбора неправильного ответа обучающийся получает за данный вопрос 0 баллов; в случае правильного ответа – 1 балл.

Оценка формируется в соответствии с критериями таблицы:

Максимальный балл	Проходной балл	Оценка
100%	не менее 50%	зачтено
менее 50%	–	не зачтено

12.2.3 Рекомендации по оцениванию в процессе наблюдения за работой на практическом занятии

Наблюдение за работой на практическом занятии – важнейший компонент контроля деятельности обучающихся и качества их обучения.

Критерии оценки в процессе наблюдения за работой на практическом занятии:

– самостоятельность выполнения задания и работы с методическими указаниями (учитывается работа в течение занятия, быстрота и способность нахождения в методических указаниях нужной информации);

– правильность выполнения работы (учитывается последовательность выполняемых действий, умение работать с инструментами и/или средствами вычислительной техники);

– рациональность использования времени, отведенного на задание (не одобряется затянутость выполнения работы с учетом индивидуальных особенностей обучающихся);

– активность обучающегося при выполнении практической работы (учитывается индивидуальная работа и работа в группе).

Критерии оценки по результатам наблюдения за работой на практическом занятии

Наименование параметра оценки	Максимальное количество баллов
Самостоятельно выполняет задания; умеет работать с методическими указаниями.	2
Указанные в методических указаниях действия выполняет в правильной последовательности; умеет работать с инструментами и/или средствами вычислительной техники.	2
Рационально использует время, отведенное на задание;	1

Наименование параметра оценки	Максимальное количество баллов
проявляет активность при выполнении практической работы.	
Итого	**5**

Максимальное количество – 5 баллов соответствует оценке «отлично», 4 балла – «хорошо», 3 балла – «удовлетворительно», 2 балла – «неудовлетворительно».

Задания и вопросы для текущего контроля знаний обучающихся по разделам (темам) учебной дисциплины приведены в приложении.

13 Промежуточная аттестация

13.1 Оценочные материалы для промежуточной аттестации обучающихся по учебной дисциплине предназначены для оценки степени достижения запланированных результатов обучения по завершению изучения учебной дисциплины в установленной учебным планом форме и позволяют определить качество усвоения изученного материала.

Итоговой формой контроля сформированности компетенций у обучающихся по учебной дисциплине является: экзамен (7 семестр); экзамен (8 семестр).

Оценочные материалы для промежуточной аттестации обучающихся включают: вопросы к экзамену; задачи к экзамену.

13.2 Оценивание обучающегося на экзамене (7 семестр)

Оценка	Требования к знаниям
отлично	Оценка «отлично» выставляется обучающемуся, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически его излагает, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, причем не затрудняется с ответом при видоизменении заданий, использует в ответе материал различной литературы, правильно обосновывает принятое нестандартное решение, обладает приемами выполнения практических задач по формированию компетенций.
хорошо	Оценка «хорошо» выставляется обучающемуся, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, обладает необходимыми умениями и приемами их выполнения.
удовлетворительно	Оценка «удовлетворительно» выставляется обучающемуся, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, испытывает сложности при

Оценка	Требования к знаниям
	выполнении практических работ и затрудняется связать теорию вопроса с практикой.
неудовлетворительно	Оценка «неудовлетворительно» выставляется обучающемуся, который не знает значительной части программного материала, неуверенно отвечает, допускает серьезные ошибки, не имеет представлений по методике выполнения практической работы. Как правило, оценка «неудовлетворительно» ставится обучающимся, которые не могут продолжить обучение без дополнительных занятий по данной учебной дисциплине.

13.3 Оценивание обучающегося на экзамене (8 семестр)

Оценка	Требования к знаниям
отлично	Оценка «отлично» выставляется обучающемуся, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически его излагает, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, причем не затрудняется с ответом при видоизменении заданий, использует в ответе материал различной литературы, правильно обосновывает принятое нестандартное решение, обладает приемами выполнения практических задач по формированию компетенций.
хорошо	Оценка «хорошо» выставляется обучающемуся, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, обладает необходимыми умениями и приемами их выполнения.
удовлетворительно	Оценка «удовлетворительно» выставляется обучающемуся, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, испытывает сложности при выполнении практических работ и затрудняется связать теорию вопроса с практикой.
неудовлетворительно	Оценка «неудовлетворительно» выставляется обучающемуся, который не знает значительной части программного материала, неуверенно отвечает, допускает серьезные ошибки, не имеет представлений по методике выполнения практической работы. Как правило, оценка «неудовлетворительно» ставится обучающимся, которые не могут продолжить обучение без дополнительных занятий по данной учебной дисциплине.

Задания и вопросы для промежуточной аттестации обучающихся приведены в приложении.

14 Учебно-методическое и информационное обеспечение учебной дисциплины

а) литература:

1 Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2024. — 490 с. — ISBN 978-5-534-00616-2. – URL: https://urait.ru/bcode/536007 (дата обращения: 13.04.2025);

2 Орлов, А. И. Прикладная статистика : практическое пособие / А.И. Орлов. – Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), 2009. – 846 с. – URL: https://biblioclub.ru/index.php?page=book&id=234537 (дата обращения: 13.04.2025);

3 Орлов, А.И. Прикладная статистика : учебное пособие / А.И. Орлов. – Москва : Национальный Открытый Университет "ИНТУИТ", 2016. – URL: https://www.studentlibrary.ru/book/intuit_263.html (дата обращения: 13.04.2025).

б) ресурсы информационно-телекоммуникационной сети «Интернет»:

1 Электронный каталог : сайт / Научно-техническая библиотека СибГИУ. – Новокузнецк, [199 –]. – URL: http://libr.sibsiu.ru. – URL: https://libr.sibsiu.ru;

2 Электронная библиотека // Научно-техническая библиотека СибГИУ : сайт. – Новокузнецк, [200 –]. – URL: http://library.sibsiu.ru/LibrELibraryFullText.asp. – Режим доступа: для авторизир. пользователей. – URL: https://library.sibsiu.ru/LibrELibraryFullText.asp;

3 Консультант студента : электронно-библиотечная система / ООО «КОНСУЛЬТАНТ СТУДЕНТА». – Москва, [200 –]. – URL: http://www.studentlibrary.ru. – Режим доступа: для авторизир. пользователей;

4 ЛАНЬ : электронно-библиотечная система : [коллекция «Единая профессиональная база знаний для технических вузов - Издательство Лань" ЭБС»] / ООО «Издательство ЛАНЬ». – Санкт-Петербург, [200 –]. – URL: http://e.lanbook.com. – Режим доступа: для авторизир. пользователей;

5 Образовательная платформа ЮРАЙТ / ООО «Электронное издательство ЮРАЙТ». – Москва, [200 –]. – URL: https://urait.ru. – Режим доступа: для авторизир. пользователей;

6 Университетская библиотека онлайн : электронно-библиотечная система / ООО «Директ-Медиа». – Москва, [200 –]. – URL: https://biblioclub.ru. – Режим доступа: для авторизир. пользователей. – URL: http://www.biblioclub.ru;

7 НАУЧНАЯ ЭЛЕКТРОННАЯ БИБЛИОТЕКА eLIBRARY.RU : база данных / ООО «НЭБ». – Москва, [200 –]. – URL: http://elibrary.ru. – Режим доступа: по подписке;

8 Электронные периодические издания ИВИС : универсальная база данных / ООО «ИВИС». – Москва, [200 –]. – URL: http://eivis.ru. – Режим доступа: по подписке.

в) лицензионное и свободно распространяемое программное обеспечение:

– 7-Zip;
– Adobe Acrobat Reader;
– Kaspersky Endpoint Security;
– Microsoft Office;
– Microsoft Windows.

г) базы данных и информационно-справочные системы:

1 ГАРАНТ : справочно-правовая система / ООО «Правовой центр «Гарант». – Кемерово, [200 –]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.;

2 КонсультантПлюс : справочно-правовая система / ООО «Информационный центр АНВИК». – Новокузнецк, [199 –]. – Режим доступа: компьютерная сеть библиотеки Сиб. гос. индустр. ун-та.;

3 Техэксперт : информационно-справочная система / ООО «Группа компаний «Кодекс». – Кемерово, [200 –]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.

15 Материально-техническое обеспечение учебной дисциплины

Материально-техническое обеспечение учебной дисциплины включает учебные аудитории, оснащенные оборудованием, компьютерной техникой, и техническими средствами обучения, в том числе:

- учебную аудиторию для проведения занятий лекционного типа, оборудованную учебной доской, экраном и мультимедийным проектором;
- учебную аудиторию для проведения занятий семинарского типа (практических занятий);
- учебную аудиторию (помещения) для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации;
- помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду, научно-техническую библиотеку СибГИУ.

Рабочая программа дисциплины составлена в соответствии с требованиями ФГОС ВО по направлению подготовки (специальности) 01.03.02 «Прикладная математика и информатика».

Составитель(и):

заведующий кафедрой Павлова Лариса Дмитриевна (кафедра прикладной математики и информатики).

Рабочая программа дисциплины рассмотрена и утверждена на заседании кафедры.

Согласована:

Заведующий научно-исследовательской лабораторией математического моделирования НИФ КемГУ, д.т.н., проф  В.О. Каледин

Приложение

Тестовые задания для входного контроля знаний обучающихся

Выберите один правильный ответ.

1. Два стрелка производят по одному выстрелу. Вероятности попадания в цель для первого и второго стрелков равны 0,7 и 0,4 соответственно. Какова вероятность того, что в цель попадет хотя бы один стрелок?

а) 0,3
б) 0,18
в) 0,28
г) 0,82

2. Игральная кость бросается один раз. Найти вероятность того, что на верхней грани выпадет *не менее двух очков, но не более пяти.*

а) 1\2
б) 2\3
в) 5\6
г) 1\3

3. В первой урне 4 черных и 6 белых шаров. Во второй урне 14 белых и 6 черных шаров. Из наудачу взятой урны вынули один шар. Найти вероятность того, что этот шар окажется белым.

а) 0,55
б) 0,45
в) 0,65
г) 0,85

4. Вероятность появления события A в 5 независимых испытаниях, проводимых по схеме Бернулли, равна 0,7. Найти вероятность того, что это событие произойдет только 2 раза

а) 0,13
б) 0,53
в) 0,19
г) 0,5

5. Вероятность того, что покупатель -мужчина сделает в магазине покупку, равна 0,1, а для женщины эта вероятность равна 0,6. Мужчины заходят в магазин в 2 раза реже, чем женщины. Какова вероятность того, что покупатель сделал покупку?

а) 0,61
б) 0,87
в) 0,43
г) 0,7

6. Найти наивероятнейшее число появлений события A, если вероятность появления этого события в каждом из 57 независимых между собой испытаниях равна 0,1.

а) 5
б) 7
в) 4
г) 6

7. В урне находятся 7 белых и 3 черных шара. Из урны вынимают 2 шара. Какова вероятность, что они оба белые?

а) 0,49
б) 0,62
в) 0,47
г) 0,55

8. Вероятность того, что после одного года обучения учебник уже нельзя будет использовать в дальнейшем, равна 0,3. Найти вероятность того, что из двух наудачу взятых учебников оба окажутся непригодными к дальнейшему использованию.

Какую формулу следует применить при решении задачи?

а) P(A+B)=P(A) + P(B) - P(AB)
б) P(AB)=P(A)P(B)
в) P(A+B)=P(A) + P(B) + P(AB)
г) P(A+B)=P(A) + P(B)

9. Как называется формула $P(H_1/A)=\dfrac{P(H_1)P(A/H_1)}{P(H_1)P(A/H_1)+P(H_2)P(A/H_2)}$

а) Формула вероятности гипотез
б) Формула полной вероятности
в) Такой формулы нет
г) Формула Бернулли

10. Найти вероятность, что событие A появится не более 14 раз в 30 независимых испытаниях, если в каждом испытании вероятность появления этого события равна 0,2

а) $\Phi(3,65)-\Phi(2,74)$
б) $2\cdot\Phi(2,74)$
в) $\Phi(3,65)+\Phi(2,74)$
г) $\Phi(2,74)-\Phi(3,65)$

Задания и вопросы для текущего контроля знаний обучающихся по разделам (темам) учебной дисциплины

1 Устный опрос на лекции

Перечень вопросов для устного опроса:

Раздел 1. Описание данных

1. Каково определение генеральной совокупности и выборки?

2. Что такое репрезентативность выборки?
3. Каково определение простой случайной выборки?
4. Что такое стратифицированная выборка?
5. Что такое мода?
6. Что такое медиана.
7. Как определяется среднее значение?
8. Какие свойства среднего?
9. Каково определение меры изменчивости данных.
10. Как выбирается мера центральной тенденции?

Раздел 2. Обработка результатов наблюдений.

1. Что такое квартили?
2. Что такое график box-plot?
3. Понятие нормального распределения.
4. Что такое правило двух и трех сигм?
5. Формулировка центральная предельной теоремы.
6. Примеры использования стандартизации.
7. Определение Т-распределения.
8. Что такое t-критерий Стьюдента?
9. Что такое дисперсионный анализ?
10 Асимметрия распределения

Раздел 3. **Временные ряды, индексы**.

1.Что такое множественное сравнение выборок?
2. Поправка Бонферрони.
3. Критерий Тьюки.
4. Множественные сравнения интерпретация результатов
5. Определение двухфакторного дисперсионного анализа.
6. Понятие корреляции.
7. Условия применения коэффициента корреляции.
8. Регрессия с одной независимой переменной.
9. Логистическая регрессия и кластерный анализ.
10. Кластерный анализ.

Раздел 4. **Гипотезы о законах распределения генеральной совокупности.**

1. Распределение Стьюдента
2. Понятие интервальной оценки, доверительного интервала. Уровень значимости, надежность оценки.
3. Понятие статистической гипотезы. Виды статистических гипотез. Процедура проверки статистических гипотез.

4. Статистический критерий. Критическая область. Уровень значимости.

5. Проверка статистических гипотез о числовом значении математического ожидания генеральной совокупности.

6. Проверка статистических гипотез о числовом значении дисперсии генеральной совокупности.

7. Проверка гипотезы о равенстве дисперсий двух генеральных распределений.

8. Критерий согласия Пирсона для проверки гипотезы нормальности распределения.

9. Точечная оценка коэффициента корреляции двух нормально распределенных случайных величин.

10. Проверка гипотезы об отсутствии корреляционной связи между двумя величинами.

Раздел 5. **Точечные и интервальные оценки параметров генеральной совокупности.**

1. Качественные статистические признаки
2. Взаимосвязи варьирующих признаков
3. Коэффициент асимметрии
4. Интервалы группировки
5. Качественно однородные группы
6. Качественные статистические признаки
7. Группировка аналитическая
8. Атрибутивные признаки группировок
9. Комплексные статистические признаки
10. Способы округления величины расчетного интервала группировки

Раздел 6. **Проверка статистических гипотез.**

1. Оптимизация проводится по конечному множеству
2. Одношаговые оценки
3. Особенности объектов нечисловой природы
4. Основная гипотеза
5. Альтернативная гипотеза
6. Статистические критерии для проверки гипотез
7. Уровень значимости
8. Ошибка первого рода
9. Ошибка второго рода
10. Область принятия гипотезы

2 Тестирование по результатам освоения лекционного материала и самостоятельной работы

Тестирование проводится для определения степени усвоения материалов лекций, а также выполнения самостоятельной работы в рамках часов, отведенных на самостоятельную работу обучающихся, в СУО «Moodle».

Перечень вопросов для тестирования:

Тестовые задания по Разделу 1

Выберите один правильный ответ.

1. Что включается в понятие «генеральная совокупность»?
а) Закон распределения
б) n независимых экспериментов
в) Эксперимент
г) подлежащая изучению статистическая совокупность
2. Что включается в понятие «случайная выборка»?
а) Закон распределения
б) n независимых экспериментов
в) Эксперимент
г) Множество из n объектов, из генеральной совокупности реально подвергшихся наблюдению
3. Репрезентативность (представительность) выборки – это
а) способность выборки воспроизводить определенные характеристики генеральной совокупности в пределах допустимых погрешностей.
б) измерение отклоняется от известного параметра генеральной совокупности больше выбранного уровня погрешности
в) результат измерения определенного параметра для данной выборки не совпадает с учетом допустимой погрешности с известным результатом измерения генеральной совокупности
г) несоответствие характеристик выборки характеристикам генеральной совокупности
4. Может ли репрезентативная выборка:
а) иметь все параметры генеральной совокупности воспроизводимых одинаково точно
б) иметь все параметры генеральной совокупности воспроизводимые по принципу дихотомии
в) одни параметры генеральной совокупности воспроизводить более точно, а другие — менее точно
г) иметь число объектов больше, чем число объектов генеральной совокупности?
5. Что такое мода?
а) величина признака, который наиболее часто встречается в данной совокупности

б) варианта, имеющая наименьшую частоту

в) периодически меняющийся признак

г) постоянный признак

6. Определение медианы:

а) в ранжированном ряду распределения одна половина ряда имеет значение признака больше половины, другая – меньше половины

б) значение признака, которое приходится на середину ранжированного ряда

в) в ранжированном ряду распределения одна треть ряда имеет значение признака больше половины, другие две трети – меньше половины

г) в ранжированном ряду распределения одна половина ряда имеет значение признака больше одной трети, другая половина – меньше двух третей

7. Как определяется среднее значение по генеральной совокупности?

а) все параметры генеральной совокупности делятся на число выборок

б) все параметры генеральной совокупности умножаются на число выборок

в) все параметры генеральной совокупности делятся на число экспериментов

г) все параметры генеральной совокупности умножаются на число экспериментов

8. Что больше? Средние: арифметическая, гармоническая, геометрическая и квадратическая?

а) x ср.гарм<= x ср.геом<= x ср.арифм<= x ср.квадр

б) x ср.гарм>= x ср.геом>= x ср.арифм>= x ср.квадр

в) x ср.арифм<= x ср.квадр <= x ср.геом <= x ср.гарм

г) x ср.арифм<= x ср.квадр<= x ср.гарм<= x ср.геом

9. Дисперсия

а) это среднее арифметическое от квадратов отклонений значений от среднего

б) это среднее гармоническое от квадратов отклонений значений от среднего

в) это среднее геометрическое от квадратов отклонений значений от среднего

г) это среднее квадратичное от квадратов отклонений значений от среднего

10. Критерий Стьюдента был введен для

а) корректировки вычислений в теории вероятности

б) корректировки вычислений при расчете ошибок экспериментальных данных

в) согласования размерностей измеряемых величин

г) согласования теоретических и экспериментальных результатов

Тестовые задания по разделу 2

Выберите один правильный ответ.

1. Вариацией значений какого-либо признака в совокупности называется:

а) различие его значений у разных единиц данной совокупности в один и тот же период или момент времени.

б) различие его значений у одинаковых единиц данной совокупности в один и тот же период или момент времени.

в) различие его значений у разных единиц данной совокупности в разные моменты времени.

г) совпадение его значений у разных единиц данной совокупности в один и тот же период или момент времени.

2. Перечень отдельных единиц совокупности в порядке возрастания (убывания) изучаемого признака называется:

а) дискретный вариационный ряд

б) ранжированный вариационный ряд

в) интервальный вариационный ряд

г) элементарный вариационный ряд

3. Дана совокупность: 123456789. квартиль первого порядка (нижний квартиль) Q1 равен:

а) 1

б) 2,5

в) 5

г) 7,5

4. Дана совокупность: 123456789. квартиль второго порядка Q2 равен:

а) 1

б) 2,5

в) 5

г) 7,5

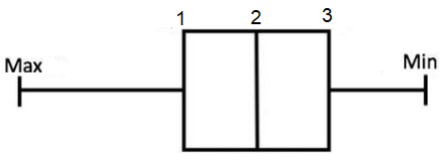
5. Дана совокупность: 123456789. квартиль третьего порядка (верхний квартиль) Q3 равен:

а) 1

б) 2,5

в) 5

г) 7,5

6. Дан график box-plot 

. Под номером 1 обозначен:

а) квартиль первого порядка (нижний квартиль) Q1

б) квартиль второго порядка Q2

в) квартиль третьего порядка (верхний квартиль) Q3

г) медиана

7. Правило трех сигм выполняется для

а) Т-распределения

б) Нормального распределения

в) Распределения Стьюдента

г) В любом распределении

8. Утверждение: «если случайная величина X представляет собой сумму очень большого числа взаимно независимых случайных величин, влияние каждой из которых на всю сумму ничтожно мало, то X имеет распределение близкое к нормальному», представляет собой формулировку.

а) критерия согласия Пирсона;

б) критерия Колмогорова;

в) критерия Тьюки;

г) центральной предельной теоремы Ляпунова;

9. Примером статистической гипотезы является следующее утверждение:

а) средняя величина определяется по формуле: $\int_a^b f(x)dx=1$

б) значение критерия, вычисленное по выборкам;

в) критерий *K* – одномерная случайная величина, все ее возможные значения принадлежат некоторому интервалу

г) генеральная совокупность распределена по закону Пуассона

10. Что такое p-уровень значимости?

а) это вероятность ошибки при отклонении нулевой гипотезы

б) ошибка корреляции

в) величина большая квартиля первого порядка

г) величина меньшая квартиля первого порядка

Тестовые задания по разделу 3

Выберите один правильный ответ.

1. К непараметрическим методам критерия множественного сравнения относят:

а) критерий Даннета

б) критерий Тьюки

в) медианный критерий

г) критерий Шеффе

2. Неравенство Бонферрони: если k-раз (k - число сравнений) применить критерий с уровнем значимости α, то вероятность хотя бы в одном случае найти различие там, где его нет, не превышает произведения k на α. Уровень значимости можно посчитать по формуле:

а) k·α

б) k/α

в) α/k

г) α - k

3. Пользуясь критерием Тьюки при наличии m групп всего возможно выполнить попарных сравнений.

а) m(m−1)/2

б) 2m(m−1)

в) m(2m−1)/2

г) m(2m−1)/4

4. Выбрать три условия применения дисперсионного анализа ANOVA

а) Наименее желателен количественный непрерывный тип данных, чем дискретные данные.

б) Независимые между собой выборки.

в) Нормальное распределение признака в статистических совокупностях, из которых извлечены выборки.

г) Независимые наблюдения в каждой из выборок.

5. В двухфакторном дисперсионном анализе исследуются факторы A и B. Сколько критериев применяются из набора:

а) Для проверки гипотезы об отсутствии эффекта фактора A

б) Для проверки гипотезы об отсутствии эффекта фактора B

в) Для проверки гипотезы об отсутствии эффекта взаимодействия факторов A и B

а) a, b

б) b, c

в) a, c

г) a, b,c

6. Корреляционная зависимость – это зависимость:

а) при которой изменение одной из величин влечет изменение среднего значения другой

б) при которой изменение одной из величин не влечет изменение среднего значения другой

в) при которой изменение одной из величин влечет изменение абсолютного значения другой

г) при которой изменение одной из величин не влечет изменение абсолютного значения другой

7. Корреляционным моментом μ_{xy} случайных величин X и Y называют

а) математическое ожидание произведения отклонений этих величин

б) математическое ожидание отношения этих величин

в) математическое ожидание отношения отклонений этих величин

г) математическое ожидание произведения этих величин

8. Коэффициент корреляции независимых случайных величин равен

а) 0

б) 1

в) 2

г) 3

9. Регрессионный анализ – это представление связи между признаками в виде функциональной зависимости, где y – результативный признак; x – факторный признак

а) $f(y)=x$

б) $y=f(x)$

в) $y=f(y/x)$

г) $x=f(y \cdot x)$

10. Цель кластерного анализа заключается:

а) в расчете коэффициентов корреляции между отдельными элементами выборки

б) в расчете коэффициентов корреляции между отдельными парными элементами выборки

в) в расчете коэффициентов регрессии между отдельными элементами выборки

г) в определении однородных в некотором смысле групп

Тестовые задания по разделу 4

Выберите один правильный ответ.

1. Среднее арифметическое является состоятельной оценкой математического ожидания

а) только если исходное распределение нормальное

б) при любом исходном распределении, если математическое ожидание существует

в) только если существует дисперсия

г) если распределение бимодально

2. Выборочная медиана может выступать оценкой

а) дисперсии

б) математического ожидания

в) коэффициента вариации

г) моды

3. ОМП для математического ожидания случайной величины, распределенной по закону Лапласа, является

а) мода

б) среднее арифметическое

в) выборочная медиана

г) дециль

4. Оптимальной оценкой математического ожидания случайной величины, распределенной по закону Лапласа, является

а) мода

б) среднее арифметическое

в) выборочная медиана

г) дециль

5. Теоретическим обоснованием перехода к пределу по двум параметрам выступает

а) теория наследования сходимости

б) центральная предельная теорема

в) закон больших чисел

г) неравенство Берри-Эссеена

6. Параметрами нормального распределения являются

а) только математическое ожидание

б) медиана и дисперсия

в) только среднеквадратическое отклонение

г) математическое ожидание и дисперсия

7. При изучении двухвыборочных статистик возникает проблема

а) перехода к пределу по объему выборки

б) перехода к пределу по двум параметрам

в) наследования сходимости

г) асимптотической размерности

8. При расчете коэффициента корреляции для учета того, что данные имеют лишь ограниченное число значащих цифр, используются

а) предельные переходы

б) неравенства Йенсена

в) поправки Шепарда

г) ранги Кендалла

9. На первом этапе решения любой прикладной задачи математическими методами/методами прикладной статистики осуществляется

а) сбор информации

б) переход от математических выводов к практической проблеме

в) внутриматематическое изучение и решение задачи

г) переход от исходной проблемы до теоретической чисто математической задачи

10. На втором этапе решения любой прикладной задачи математическими методами/методами прикладной статистики осуществляется

а) сбор информации

б) переход от математических выводов к практической проблеме

в) внутриматематическое изучение и решение задачи

г) переход от исходной проблемы до теоретической чисто математической задачи

Тестовые задания по разделу 5

Выберите один правильный ответ.

1. Аналитическая группировка применяется для…
а) характеристики структурных сдвигов
б) характеристики структуры совокупности
в) характеристики взаимосвязи между изучаемыми явлениями и их признаками
г) разделения совокупности на качественно однородные типы единиц
2. Величина интервала для группировки с неравными интервалами определяется по
а) арифметической прогрессии
б) геометрической прогрессии
в) экспоненциональмому закону
г) логарифмическому закону
3. В результате округления величины расчетного интервала группировки равного 167 получим …
а) 150
б) 170
в) 200
г) 160
г) Группировка, в которой исследуются взаимосвязи варьирующих признаков в пределах однородной совокупности, называется
а) комбинированной
б) типологической
в) аналитической
г) структурной
5. Группировка, в которой качественно неоднородная совокупность делится на отдельные качественно однородные группы и на этой

основе выявляются экономические типы явлений, называется ...
а) структурной
б) комбинированной
в) типологической
г) аналитической
6. Если мода, медиана и средняя арифметическая статистического ряда совпадают, то коэффициент асимметрии
а) больше единицы
б) равен 10%
в) равен единице
г) меньше единицы
7. Интервалы группировки могут быть …
а) фиксированными

б) закрытыми
в) связанными
г) открытыми
8. Исследование взаимосвязей варьирующих признаков в пределах однородной совокупности называется ….. группировкой.
а) структурной
б) типологической
в) множественной
г) аналитической
9. К атрибутивным признакам группировок относят …
а) пол
б) среднедушевой доход
в) уровень образования
г) возраст
10. Качественные статистические признаки подразделяются на
а) агрегатные
б) комплексные
в) альтернативные
г) порядковые

Тестовые задания по разделу 6

Выберите один правильный ответ.

1. На третьем, последнем этапе решения любой прикладной задачи математическими методами/методами прикладной статистики осуществляется
а) сбор информации
б) переход от математических выводов к практической проблеме
в) внутриматематическое изучение и решение задачи
г) переход от исходной проблемы до теоретической чисто математической задачи
2. Одна математическая модель может применяться для решения
а) только одной задачи
б) ряда сходных по прикладной сущности задач
в) разных по прикладной сущности задач
3. Если оптимизация проводится по конечному множеству, то оптимум
а) всегда достигается и единственен
б) всегда достигается, но может быть неединственным
в) всегда достигается, но обязательно будет неединственным
г) не достигается
4. Метод Монте-Карло - это
а) абстракция
б) метод статистических испытаний

в) метод складного ножа

г) метод предельных переходов

5. Верно, что

а) одношаговые оценки обладают худшими асимтотическими свойствами, чем ОМП

б) с вычислительной точки зрения одношаговые оценки проще ОМП

в) при малых объемах выборок одношаговые оценки совпадают с ОМП

6. В качестве средних величин для данных, измеренных в порядковой шкале, рекомендуется использовать порядковые статистика, например

а) моду

б) среднее арифметическое

в) медиану

г) люсиан

7. Государственный комитет по статистике ранее назывался

а) Комитетом по математической статистике и выборочным обследованиям

б) Центральным статистическим управлением

в) Управлением по делам статистики

8. Метрика Хаусдорфа - это одно из расстояний между

а) множествами

б) ранжировками

в) полутолерантностями

г) функциями распределения

9. Верно, что в пространствах объектов нечисловой природы

а) нельзя ввести метрику

б) не решаются оптимизационные задачи

в) нет операции сложения

10. При изучении предпочтений потребителей решение вопроса о том, различаются ли мнения двух групп потребителей, сводится к

а) применению кластерного анализа

б) проверке той или иной статистической гипотезы однородности

в) введению специальной метрики в пространстве нечисловой природы

г) применению регрессионного анализа

3 Выполнение практических работ

Задания для практических работ, теоретический и методический материал для самостоятельного изучения и вопросы для собеседования представлены в методических указаниях к практическим занятиям.

Примеры заданий для практической работы:

Раздел 1. Характеристики двумерной случайной величины

Дана плотность совместного распределения двумерной случайной величины (X,У):

$$f(x,y)=\begin{cases}0 & при \quad x<0;\\ Cxe^{-yx} & при \quad x\geq 0.\end{cases}$$

Найти:

1) Функцию распределения системы двух случайных величин ($F(x,y)$).
2) Функции распределения величин ($f_1(x)$,$f_2(y)$).
3) Математическое ожидание случайных величин (MX, MY).
4) Дисперсию случайных величин (DX, DY).
5) Вероятность $p(X,Y)$.

Раздел 2. Парный корреляционный анализ

По данным X и Y в таблице, соответствующим номеру варианта, требуется:

1. Вычислить выборочный коэффициент корреляции *r* между X и Y.
2. Проверить значимость истинного коэффициента корреляции ρ с помощью таблицы Фишера – Йейтса .
3. В случае, если ρ значим, найти для него интервальную оценку с помощью *z* - преобразования Фишера.

Номер варианта	Признаки	Номер варианта	Признаки	Номер варианта	Признаки
1	Y_1X_1	10	Y_2X_1	19	Y_3X_1
2	Y_1X_2	11	Y_2X_2	20	Y_3X_2
3	Y_1X_3	12	Y_2X_3	21	Y_3X_3
4	Y_1X_4	13	Y_2X_4	22	Y_3X_4
5	Y_1X_5	14	Y_2X_5	23	Y_3X_5
6	Y_1X_6	15	Y_2X_6	24	Y_3X_6
7	Y_1X_7	16	Y_2X_7	25	Y_3X_7
8	Y_1X_8	17	Y_2X_8	26	Y_3X_8
9	Y_1X_9	18	Y_2X_9	27	Y_3X_9

Обозначения и наименования признаков:

Y_1 - производительность труда, млн. руб./ чел.;

Y_2 - индекс снижения себестоимости продукции;

Y_3 - рентабельность, %;

X_1 - трудоемкость единицы продукции, чел. - час.;

X_2 - удельный вес рабочих в составе промышленно-производственного персонала;

X_3 - удельный вес покупных изделий;

X_4 - коэффициент сменности оборудования, смен;

X_5 - премии и вознаграждения на одного работника промышленно-производственного персонала, млн. руб.;

X_6 - удельный вес потерь от брака, %;

X_7 - фондоотдача активной части основных производственных фондов, руб./ руб.;

X_8 - среднегодовая стоимость основных производственных фондов, млрд. руб.;

X_9 - фондовооруженность труда, млн. руб./ чел.

Значения признаков заданы в таблице.

№	Y_1	Y_2	Y_3	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
1	9,3	62,0	10,6	0,23	0,62	0,40	1,35	0,88	0,15	1,91	39,53	5,35
2	9,8	53,1	9,1	0,43	0,76	0,19	1,39	0,57	0,34	1,68	40,41	3,90
3	9,1	56,5	23,4	0,26	0,71	0,44	1,27	1,70	0,09	1,89	37,02	4,88
4	5,4	30,1	9,7	0,45	0,74	0,25	1,10	0,84	0,05	1,02	41,08	5,65
5	6,6	18,1	9,1	0,38	0,72	0,02	1,23	1,04	0,48	0,88	42,39	8,85
6	4,3	13,6	5,4	0,42	0,68	0,06	1,39	0,66	0,41	0,62	37,39	8,52
7	7.4	89,8	9,9	0,30	0,77	0,15	1,38	0,86	0,62	1,09	71,78	7,19
8	6,6	76,6	19,1	0,37	0,77	0,24	1,35	1,27	0,50	1,32	81,32	5,38
9	5,5	32,3	6,6	0,34	0,72	0,11	1,24	0,68	1,20	0,68	59,92	9,27
10	9,4	99,6	14,2	0,23	0,79	0,47	1,40	0,86	0,21	2,30	97,34	4,36
11	5,7	90,8	8,0	0,41	0,71	0,20	1,28	0,45	0,66	1,43	80,83	4,16
12	5,2	82,1	17,5	0,46	0,79	0,24	1,33	0,74	0,74	1,82	59,42	3,13
13	10,0	76,2	17,2	0,22	0,76	0,54	1,22	1,03	0,32	2,62	39,96	4,02
14	6,7	37,1	12,9	0,31	0,78	0,29	1,36	0,96	0,39	1,24	37,21	5,82
15	9,4	51,6	13,2	0,24	0,70	0,56	1,20	0,98	0,28	2,03	32,87	5,01

Раздел 3. Множественный корреляционный анализ

По данным *n*=10 машиностроительных предприятий в таблице методами корреляционного анализа исследовать взаимосвязь между следующими показателями: X_1 – рентабельность (%); X_2 – премии и вознаграждения на одного работника (млн. руб.); X_3 – фондоотдача.

№	X_1	X_2	X_3
1	12,26	1,33	1,55
2	11,16	1,14	1,40
3	12,72	1,70	1,77
4	11,82	0,33	1,35
5	1163	0,48	1,81
6	8,12	0,67	1,58
7	24,83	1,52	1,34
8	22,39	1,60	1,69
9	15,68	0,74	1,84
10	11,05	0,60	2,16

Требуется:

а) рассчитать векторы средних и среднеквадратических отклонений, матрицу парных коэффициентов корреляции *(* $\overline{X}$*, S, R);*

б) проверить при α=0,05 значимость парного коэффициента корреляции ρ_{12}и найти его интервальную оценку с доверительной вероятностью γ=0,95;

в) по корреляционной матрице R рассчитать частный коэффициент корреляции $r_{12/3}$;

г) проверить при α=0,05 значимость частного коэффициента корреляции $\rho_{12/3}$ и определить его интервальную оценку при γ=0,95;

д) по корреляционной матрице R вычислить оценку множественного коэффициента корреляции $r_{1/23}$ и при α=0,05 проверить гипотезу H_0: $r_{1/23}$=0.

Задание выполняется по вариантам. Каждый должен вычеркнуть объект № *k*, соответствующий номеру варианта. Так, например, если номер варианта 2, то нужно вычеркнуть второй объект.

Раздел 4. Парный регрессионный анализ

По территориям региона приводятся данные за год (таблица, в которой 1 *p* – число букв в полном имени, 2 *p* – число букв в фамилии).

Номер региона	Среднедушевой прожиточный минимум в день одного трудоспособного, руб., *x*	Среднедневная заработная плата, руб., *y*
1	77+ *p*	137+ 2 *p*
2	81+ 2 *p*	147
3	86	134+ *p*
4	78	156
5	105	158+ *p*
6	104+ *p*	191
7	67	136
8	90	159+ 2 *p*
9	73+ 2 *p*	152
10	87	162
11	86	145+ 2 *p*
12	111+ *p*	

Требуется:

1. Построить линейное уравнение парной регрессии *y* по *x* .

2. Рассчитать линейный коэффициент парной корреляции, коэффициент детерминации и среднюю ошибку аппроксимации.

3. Оценить статистическую значимость уравнения регрессии в целом и отдельных параметров регрессии и корреляции с помощью *F* - критерия Фишера и *t* -критерия Стьюдента.

4. Выполнить прогноз заработной платы *y* при прогнозном значении среднедушевого прожиточного минимума *x* , составляющем 107% от среднего уровня.

5. Оценить точность прогноза, рассчитав ошибку прогноза и его доверительный интервал.

6. На одном графике отложить исходные данные и теоретическую прямую.

7. Проверить вычисления в MS Excel.

Раздел 5. Множественный регрессионный анализ

По 20 предприятиям региона изучить зависимость выработки продукции на одного работника *y* (тыс. руб.) от ввода в действие новых основных фондов 1 *x* (% от стоимости фондов на конец года) и от удельного веса рабочих высокой квалификации в общей численности рабочих 2 *x* (%) (таблица, в которой 1 *p* – число букв в полном имени, 2 *p* – число букв в фамилии).

Номер предприятия	*y*	x_1	x_2
1	7,1	3,6+0.1*p*	11,0
2	7,2	3,8	13,0
3	7,0	3,9	15,0
4	7,3	4,0	17,0
5	7,4	3,8+0,1*p*	18,0
6	8,0	4,8	19,0
7	8,0	5,3	19,0
8	8,0	5,4	20,0
9	8,0	5,6–0,1*p*	20,0
10	10,0	6,8	21,0
11	9,0	6,0+0,2*p*	21,0
12	11,0	6,4	22,0
13	9,0	6,9	22,0
14	11,0	7,2	25,0
15	12,0	8,0–0,2*p*	28,0
16	12,0	8,2	29,0
17	12,0	8,1	30,0
18	12,0	8,6	31,0
19	14,0	9,6	32,0
20	14,0	9,0+,2*p*	36,0

Требуется:

1. Построить линейную модель множественной регрессии. Записать стандартизованное уравнение множественной регрессии. На основе стандартизованных коэффициентов регрессии и средних

коэффициентов эластичности ранжировать факторы по степени их влияния на результат.

2. Найти коэффициенты парной, частной и множественной корреляции. Проанализировать их.

3. Найти скорректированный коэффициент множественной детерминации. Сравнить его с нескорректированным (общим) коэффициентом детерминации.

4. С помощью *F* -критерия Фишера оценить статистическую надежность уравнения регрессии и коэффициента детерминации

5. С помощью *t* -критерия Стьюдента оценить статистическую значимость параметров чистой регрессии.

6. С помощью частных *F* -критериев Фишера оценить целесообразность включения в уравнение множественной регрессии фактора x_1после x_2и фактора x_2после *x1* .

7. Составить уравнение линейной парной регрессии, оставив лишь один значащий фактор.

8. Проверить вычисления в MS Excel.

Раздел 6. Многомерный статистический анализ

Исходные данные: данные РМЭЗ (индивидуальные данные) за 2007г.

Описание проекта РМЭЗ: http://hse.ru/science/rlms.

Сформулируйте 3 задачи однофакторного дисперсионного анализа по данным РМЭЗ.

Факторы (*X*) должны быть:

количественные (интервальные)

- порядковые
- номинальные

Определите переменные *X* и *Y* (кратко опишите дескриптивные характеристики показателей).

В задачах:

- сформулируйте гипотезу дисперсионного анализа.
- при α=0,05 проверьте с помощью процедур SPSS существенность влияния *X* на *Y*;

В отчет представить результаты расчетов в *MS Excel*, описать подробно, чему равны суммы квадратов: Q_A, $Q_{ост}$, $Q_{общ}$, $F_{набл}$ (сравнить с $F_{крит}$) и сделать выводы.

4 Вопросы к экзамену для промежуточной аттестации (7 семестр)

1. Понятие генеральной совокупности и выборки, репрезентативность выборки. Простая случайная выборка. Стратифицированная выборка. Групповая выборка.

2. Представление выборки. (Таблица частот и интервальная таблица частот. Графическое представление выборки);

3. Выборочное среднее, мода, медиана. Квартили, декатили, персентили.

4. Измерение разброса: размах, выборочная дисперсия, выборочное среднее квадратическое отклонение (стандартное отклонение)

5. Коэффициент вариации. О симметричных и несимметричных распределениях.

6. Вычисление выборочного среднего и выборочной дисперсии для объединения двух выборок. Общая, межгрупповая и внутригрупповая дисперсии.

7. Кривая Лоренца и показатели концентрации.

8. Двумерные выборки. Графическое представление двумерных выборок. Диаграммы рассеяния.

9. Выборочный коэффициент корреляции — числовая характеристика двумерной выборки.

10. Метод наименьших квадратов;

11. Уравнения регрессии.;

12. Расчет коэффициентов линейного уравнения регрессии по сгруппированным данным;

13. Индекс корреляции.

14. Тренды временных рядов. Линейный тренд. Параболический тренд. Показательная функция. Исключение трендовой составляющей.

15. Скользящие средние. Сезонные колебания и индексы сезонности;

16. Индивидуальные (частные) индексы. Общие индексы.

17. Агрегатные индексы. Средние индексы.

18. Индексы цен. Дефлятирование стоимостных величин.

19. (Проверка гипотезы о законе распределения генеральной совокупности по критерию Пирсона.

20. Проверка гипотезы о законе распределения Пуассона.

5 Задачи к экзамену для промежуточной аттестации (7 семестр)

1. Известны данные о числе посещений пункта семейной медицины за пять дней с начала месяца: 1 апреля – 15 чел.; 2 апреля – 16 чел.; 3 апреля – 19 чел.; 4 апреля – 13 чел.; 5 апреля – 15 чел. Найдите тип динамического ряда.

2. План на отчетный период предусматривал снижение издержек производства на предприятии на 5%. По итогам отчетного периода достигнуто снижение издержек производства по сравнению с прошлым периодом на 6%. Какова относительная величина планового задания.

3. В регионе на 1000 чел. в трудоспособном возрасте приходится 202 чел. моложе трудоспособного возраста и 374 чел. старше трудоспособного возраста. Чему равен удельный вес населения трудоспособного возраста в регионе?

4. Численность населения в регионе А составляет 1620 тыс. чел. и при этом в среднем за год сокращается на 24 тыс. чел. Численность населения региона Б составляет 1480 тыс. чел. и растет в среднем на 20 тыс. чел. в год. Через сколько лет регионы сравняются по численности населения при условии, что демографические тенденции в регионах останутся неизменными:

5. В регионе за отчетный период цены на продовольственные и непродовольственные товары в розничной торговле выросли на 9 и 11% соответственно. Определите, как в среднем изменились цены розничной торговли, если известно, что на долю продовольственных товаров приходится 35% розничного товарооборота

6. Изучается связь между типом телефонных переговоров (внутригородские, междугородние и международные) и продолжительностью телефонных разговоров. Получены следующие значения дисперсии: общая – 620, межгрупповая – 310, средняя из внутригрупповых – 310. Какого вида связь между типом переговоров и продолжительностью разговора?

7. По данным случайного опроса 400 избирателей в регионе А и 400 избирателей в регионе Б о намерении участвовать в выборах губернатора, в регионе А заявили 25%, а в регионе Б – 30% опрошенных. Существенны или случайны выявленные различия в намерениях избирателей регионов А и Б?

8. Динамика суммы прибыли (в млн руб.) за ряд лет описывается уравнением $Y = 25,7 \cdot 1,15^t$, где t = 1,2… . На сколько процентов ежегодно возрастает прибыль?

9. В библиотеке изучаются сроки пользования читателей книгами. С этой целью из 20000 выданных книг в порядке бесповторной выборки обследовано 5% их числа. Установлено, что средний срок пользования книгой составил 18 дней при коэффициенте вариации 67%. Количество книг, находящихся в пользовании более установленного срока, оказалось равным 196. С вероятностью 0,93 определите предельные значения среднего срока пользования книгой и количество книг, возвращенных с нарушением установленного срока пользования.

10. В результате социологического опроса 100 домохозяйств области получены следующие характеристики формы распределения по среднедушевому доходу в месяц: • коэффициент асимметрии (–0,547); • показатель эксцесса (–0,91). Оцените, существенно ли отличается полученное распределение от нормального.

6 Вопросы к экзамену для промежуточной аттестации (8 семестр)

1. Генеральная совокупность и выборка.
2. Переменные. Типы переменных.
3. Основные понятия описательной статистики.
4. Меры изменчивости.
5. Квартили распределения и график box-plot.
6. Нормальное распределение.
7. Распределение “хи-квадрат”
8. Распределение Стьюдента
9. Понятие интервальной оценки, доверительного интервала. Уровень значимости, надежность оценки.
10. Понятие статистической гипотезы. Виды статистических гипотез. Процедура проверки статистических гипотез.
11. Статистический критерий. Критическая область. Уровень значимости.
12. Проверка статистических гипотез о числовом значении математического ожидания генеральной совокупности.
13. Проверка статистических гипотез о числовом значении дисперсии генеральной совокупности.
14. Проверка гипотезы о равенстве дисперсий двух генеральных распределений.
15. Критерий согласия Пирсона для проверки гипотезы нормальности распределения.
16. Точечная оценка коэффициента корреляции двух нормально распределенных случайных величин.
17. Проверка гипотезы об отсутствии корреляционной связи между двумя величинами.
18. Простая регрессия. Выборочная оценка коэффициентов регрессии.
19. Множественная регрессия.
20. Кластерный анализ.

7 Задачи к экзамену для промежуточной аттестации (8 семестр)

1. Линейные коэффициенты корреляции между процентом выполнения норм выработки (y), квалификацией (x) и стажем работы рабочих предприятия (z) оказались равными: r_{yx} = +0,587; r_{yz} = +0,348; r_{xz} = 0,108. Определите совокупный коэффициент корреляции между процентом выполнения норм выработки и двумя определяющими его факторами (квалификацией и стажем работы).
2. Линейные коэффициенты корреляции между процентом выполнения норм выработки (y), квалификацией (x) и стажем работы рабочих предприятия (z) оказались равными: r_{yx} = +0,587; r_{yz} = +0,348; r_{xz} = 0,108. Определите частные коэффициенты корреляции между: а)

процентом выполнения норм выработки и квалификацией рабочих; б) процентом выполнения норм выработки и стажем работы рабочих.

3. Дисперсия годовых расходов на приобретение продуктов питания составляет в регионе А 16800, в регионе Б – 24100. В каком из регионов при равных объемах выборки ошибка выборки будет больше?

4. При анализе изменения структуры иностранных инвестиций в двух регионах получены следующие значения коэффициента корреляции рангов по видам инвестиций: регион А – 0,2, регион Б – 0,8 . В каком из регионов изменения структуры были более существенными?

5. Во сколько раз изменится коэффициент корреляции, определяющий зависимость между переменными y и x1 при включении в анализ дополнительной переменной x2, если известно, что r_{yx1} = = 0,897, а r_{x1x2} = –0,425?

6. Текущие затраты на охрану окружающей среды на предприятии в сопоставимых ценах выросли за 6 месяцев с начала года на 18%. Во сколько раз изменится среднемесячный темп роста затрат?

7. Если таблица сопряженности для двух переменных имеет 5 строк и 6 столбцов. Чему равно число степеней свободы для статистики хи-квадрат?

8. Имеются следующие выборочные данные о денежных

Группа вкладчиков	Число вкладов	Средняя сумма вклада, тыс. руб.	Коэффициент вариации суммы вклада, %
Горожане	70	34	20
Сельские жители	30	16	30

вкладах населения в коммерческие банки района. Определите тесноту связи между средним размером вклада и группой вкладчиков.

9. По 50 предприятиям отрасли изучается зависимость между рентабельностью производства (прибыль, % к стоимости основных и оборотных фондов), производительностью труда (x_1 – млн руб. на одного работника) и средним возрастом производственного оборудования (x_2 – лет). Получено следующее уравнение регрессии:

$$y = 6,9 + \underset{(0,1)}{0,67} \cdot x_1 - \underset{(0,125)}{0,3} \cdot x_2.$$

. В скобках приведены стандартные ошибки параметров уравнения. R^2=0,98. Оцените статистическую значимость коэффициентов регрессии и уравнения регрессии в целом.

10. Проводится выборочная проверка качества выпускаемых электроламп. Случайно отобраны 1000 ламп, что составляет 4% общего выпуска. Установлено, что средний срок горения лампы составляет 2300 ч, а коэффициент вариации срока горения – 30%. С какой вероятностью можно утверждать, что средний срок горения всех изготовленных электроламп будет не менее 2256 и не более 2344 ч?