



**САМАРСКИЙ
ПОЛИТЕХ**
Опорный университет

МИНОБРНАУКИ РОССИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Самарский государственный технический университет»
(ФГБОУ ВО «СамГТУ»)



УТВЕРЖДАЮ:

Первый проректор-проректор по
учебной работе

/ Д.Е. Овчинников

25 09 2026 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.О.03.02 «Математическая статистика»

Код и направление подготовки (специальность)	02.03.03 Математическое обеспечение и администрирование информационных систем
Направленность (профиль)	Анализ данных в бизнес-аналитике
Квалификация	Бакалавр
Форма обучения	Очная
Год начала подготовки	2026
Институт / факультет	Институт инженерно-экономического и гуманитарного образования
Выпускающая кафедра	кафедра "Высшая математика"
Кафедра-разработчик	кафедра "Высшая математика"
Объем дисциплины, ч. / з.е.	108 / 3
Форма контроля (промежуточная аттестация)	Экзамен

Содержание

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы	4
2. Место дисциплины (модуля) в структуре образовательной программы	5
3. Объем дисциплины (модуля) в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся	5
4. Содержание дисциплины (модуля), структурированное по темам (разделам), с указанием отведенного на них количества академических часов и видов учебных занятий	6
4.1 Содержание лекционных занятий	6
4.2 Содержание лабораторных занятий	8
4.3 Содержание практических занятий	9
4.4. Содержание самостоятельной работы	11
5. Перечень учебной литературы и учебно-методического обеспечения по дисциплине (модулю)	11
6. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения	12
7. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», профессиональных баз данных, информационно-справочных систем	12
8. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)	12
9. Методические материалы	13
10. Фонд оценочных средств по дисциплине (модулю)	15

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

Наименование категории (группы) компетенций	Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Результаты обучения (знать, уметь, владеть, соотнесенные с индикаторами достижения компетенции)
Общепрофессиональные компетенции			
Теоретические и практические основы профессиональной деятельности	ОПК-1 Способен применять фундаментальные знания, полученные в области математических и (или) естественных наук, и использовать их в профессиональной деятельности	ОПК-1.1 Знать: алгебраические, теоретико-вероятностные и статистические методы для использования в анализе больших данных	Знать алгебраические, теоретико-вероятностные и статистические методы для использования в анализе больших данных.
		ОПК-1.2 Уметь: строить математические модели, отражающие закономерности и взаимосвязи между факторами больших данных	Уметь строить математические модели, отражающие закономерности и взаимосвязи между факторами больших данных.
		ОПК-1.3 Владеть: алгебраическими, теоретико-вероятностными и статистическими понятиями для построения математических моделей	Владеть алгебраическими, теоретико-вероятностными и статистическими понятиями для построения математических моделей.
	ОПК-2 Способен применять современный математический аппарат, связанный с проектированием, разработкой, реализацией и оценкой качества программных продуктов и программных комплексов в различных областях человеческой деятельности	ОПК-2.1 Знать: вычислительные алгоритмы оптимизации и численные методы решения больших линейных систем	Знать вычислительные алгоритмы оптимизации (градиентный спуск, ЕМ-алгоритм) и численные методы решения больших линейных систем (метод наименьших квадратов, сингулярное разложение).

		ОПК-2.2 Уметь: реализовывать различные математические модели (в том числе регрессионные) в виде вычислительных алгоритмов	Уметь реализовывать различные математические модели (в том числе регрессионные) в виде вычислительных алгоритмов с использованием библиотек Python/NumPy/SciPy.
		ОПК-2.3 Владеть: навыками скользящего контроля алгоритма, настроенного на обучающей выборке, и навыками регуляризации алгоритма для предотвращения его переобучения	Владеть навыками скользящего контроля (cross-validation) алгоритма, настроенного на обучающей выборке, и навыками регуляризации алгоритма (L1/L2, Ridge, Lasso) для предотвращения его переобучения.

2. Место дисциплины (модуля) в структуре образовательной программы

Место дисциплины (модуля) в структуре образовательной программы: **обязательная часть**

Код компетенции	Предшествующие дисциплины	Параллельно осваиваемые дисциплины	Последующие дисциплины
ОПК-1	Алгебра; Дифференциальные уравнения; Математический анализ; Теория вероятностей; Теория графов; Теория функций комплексной переменной	Случайные процессы	Подготовка к процедуре защиты и защита выпускной квалификационной работы; Уравнения математической физики
ОПК-2	Вычислительная линейная алгебра; Дискретная математика; Методы оптимизации и исследования операций; Теория вероятностей; Теория графов	Вычислительная математика; Случайные процессы	Подготовка к процедуре защиты и защита выпускной квалификационной работы

3. Объем дисциплины (модуля) в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся

Вид учебной работы	Всего часов / часов в электронной форме	5 семестр часов / часов в электронной форме
Аудиторная контактная работа (всего), в том числе:	64	64
Лабораторные работы	16	16
Лекции	32	32
Практические занятия	16	16
Внеаудиторная контактная работа, КСР	4	4

Самостоятельная работа (всего), в том числе:	4	4
подготовка к лабораторным работам	2	2
подготовка к практическим занятиям	2	2
Контроль	36	36
Итого: час	108	108
Итого: з.е.	3	3

4. Содержание дисциплины (модуля), структурированное по темам (разделам), с указанием отведенного на них количества академических часов и видов учебных занятий

№ раздела	Наименование раздела дисциплины	Виды учебной нагрузки и их трудоемкость, часы				
		ЛЗ	ЛР	ПЗ	СРС	Всего часов
1	Математические основы статистического вывода	12	6	4	1	23
2	Параметрические и непараметрические методы проверки гипотез	6	4	4	1	15
3	Регрессионный анализ и моделирование зависимостей	10	6	4	1	21
4	Вычислительная статистика, регуляризация и снижение размерности	4	0	4	1	9
	КСР	0	0	0	0	4
	Контроль	0	0	0	0	36
	Итого	32	16	16	4	108

4.1 Содержание лекционных занятий

№ занятия	Наименование раздела	Тема лекции	Содержание лекции (перечень дидактических единиц: рассматриваемых подтем, вопросов)	Количество часов / часов в электронной форме
5 семестр				
1	Математические основы статистического вывода	Математические основы статистического вывода	Выборочный метод. Вариационные ряды, эмпирическая функция распределения. Теорема Гливенко-Кантелли.	2
2	Математические основы статистического вывода	Математические основы статистического вывода	Статистики выборки и их распределения. Распределения Пирсона, Стьюдента, Фишера-Снедекора.	2
3	Математические основы статистического вывода	Математические основы статистического вывода	Точечное оценивание параметров. Свойства оценок (несмещенность, состоятельность, эффективность). Неравенство Рао-Крамера.	2

4	Математические основы статистического вывода	Математические основы статистического вывода	Методы получения оценок: метод моментов, метод максимального правдоподобия (ММП). Численные аспекты оптимизации функции правдоподобия.	2
5	Математические основы статистического вывода	Математические основы статистического вывода	Интервальное оценивание. Доверительные интервалы для математического ожидания и дисперсии при известной и неизвестной дисперсии генеральной совокупности.	2
6	Математические основы статистического вывода	Математические основы статистического вывода	Введение в проверку гипотез. Общая схема проверки. Ошибки I и II рода. Мощность критерия. P-value.	2
7	Параметрические и непараметрические методы проверки гипотез	Параметрические и непараметрические методы проверки гипотез	Параметрические критерии значимости. Z-критерий для пропорций. t-критерий Стьюдента (для одной и двух независимых/зависимых выборок). F-критерий равенства дисперсий.	2
8	Параметрические и непараметрические методы проверки гипотез	Параметрические и непараметрические методы проверки гипотез	Непараметрические критерии согласия: критерий Колмогорова-Смирнова, критерий Пирсона для проверки теоретического закона распределения.	2
9	Параметрические и непараметрические методы проверки гипотез	Параметрические и непараметрические методы проверки гипотез	Непараметрические критерии сдвига и однородности: критерий знаков, Вилкоксона-Манна-Уитни, критерий знаковых рангов Уилкоксона. Применение A/B-тестирования в бизнесе.	2
10	Регрессионный анализ и моделирование зависимостей	Регрессионный анализ и моделирование зависимостей	Модель парной линейной регрессии. Метод наименьших квадратов (МНК). Геометрическая интерпретация проекции вектора отклика на пространство признаков.	2
11	Регрессионный анализ и моделирование зависимостей	Регрессионный анализ и моделирование зависимостей	Матричная форма множественной линейной регрессии. Нормальные уравнения. Предпосылки теоремы Гаусса-Маркова.	2
12	Регрессионный анализ и моделирование зависимостей	Регрессионный анализ и моделирование зависимостей	Оценка качества регрессионной модели. Коэффициент детерминации и скорректированный коэффициент детерминации. Дисперсионный анализ (ANOVA) регрессии. Проверка адекватности модели (анализ остатков).	2
13	Регрессионный анализ и моделирование зависимостей	Регрессионный анализ и моделирование зависимостей	Обобщенная линейная модель. Логистическая регрессия как базовый алгоритм классификации. Функция потерь кросс-энтропии.	2
14	Регрессионный анализ и моделирование зависимостей	Регрессионный анализ и моделирование зависимостей	Проблема мультиколлинеарности факторов. Численные методы решения больших систем нормальных уравнений (SVD-разложение, QR-разложение).	2

15	Вычислительная статистика, регуляризация и снижение размерности	Вычислительная статистика, регуляризация и снижение размерности	Компромисс смещения и дисперсии (Bias-Variance Tradeoff). Переобучение алгоритмов машинного обучения.	2
16	Вычислительная статистика, регуляризация и снижение размерности	Вычислительная статистика, регуляризация и снижение размерности	Валидация и оценка обобщающей способности. Скользящий контроль (K-Fold Cross-Validation). Bootstrap-методы оценки стабильности коэффициентов.	2
Итого за семестр:				32
Итого:				32

4.2 Содержание лабораторных занятий

№ занятия	Наименование раздела	Тема лабораторного занятия	Содержание лабораторного занятия (перечень дидактических единиц: рассматриваемых подтем, вопросов)	Количество часов / часов в электронной форме
5 семестр				
1	Математические основы статистического вывода	Оценка параметров распределений и проверка статистик	Генерация псевдослучайных величин методом обратной функции и Бокса-Мюллера. Построение гистограмм, эмпирической функции распределения.	2
2	Математические основы статистического вывода	Метод максимального правдоподобия (ММП).	Реализация ММП «с нуля» через градиентный спуск для оценки параметра экспоненциального распределения или дисперсии нормального распределения. Сравнение полученной оценки с аналитическим решением метода моментов. Визуализация функции правдоподобия.	2
3	Параметрические и непараметрические методы проверки гипотез	Параметрические критерии проверки гипотез	Программная реализация Z-теста для пропорций и t-критерия Стьюдента для двух независимых выборок без использования готовых функций (ttest_ind). Применение критериев к бизнес-данным (например, сравнение конверсии двух лендингов A/B-теста). Использование F-критерия для сравнения волатильности активов.	2
4	Параметрические и непараметрические методы проверки гипотез	Непараметрические критерии согласия и сдвига.	Применение критерия Колмогорова-Смирнова и Пирсона для проверки гипотезы о нормальном законе распределения остатков модели. Реализация U-критерия Манна-Уитни для анализа несимметричных данных (например, распределение чеков в ритейле).	2

5	Регрессионный анализ и моделирование зависимостей	Классический метод наименьших квадратов (МНК).	Построение парной линейной регрессии. Реализация матричного МНК через нормальное уравнение средствами numpy.linalg. Диагностика остатков: построение графиков Residuals?Fitted, Q-Q plot, Scale-Location. Выявление гетероскедастичности тестом Голдфелда-Квандта.	2
6	Регрессионный анализ и моделирование зависимостей	Множественная линейная регрессия и диагностика.	Работа с многомерными данными. Расчет VIF (фактора инфляции дисперсии) для выявления мультиколлинеарности вручную (через вспомогательные регрессии признаков друг на друга). Оценка качества модели. Пошаговый отбор признаков	2
7	Регрессионный анализ и моделирование зависимостей	Логистическая регрессия и регуляризация Ridge/Lasso.	Решение задачи бинарной классификации (например, скоринг заемщиков). Реализация логистической регрессии через минимизацию кросс-энтропии. Настройка гиперпараметров регуляризации методами Grid Search CV и Randomized Search CV. Анализ изменения весов признаков при увеличении силы регуляризации. Сравнение моделей Lasso (отбор признаков) и Ridge (снижение дисперсии).	2
8	Математические основы статистического вывода	Оценка параметров распределений и проверка статистик.	Проверка свойств оценок (несмещенность, состоятельность) на симуляциях Монте-Карло. Расчет доверительных интервалов бутстреп-методом для медианы.	2
Итого за семестр:				16
Итого:				16

4.3 Содержание практических занятий

№ занятия	Наименование раздела	Тема практического занятия	Содержание практического занятия (перечень дидактических единиц: рассматриваемых подтем, вопросов)	Количество часов / часов в электронной форме
5 семестр				
1	Математические основы статистического вывода	Вариационные ряды и точечные оценки.	Построение дискретных и интервальных вариационных рядов. Расчет выборочных характеристик: среднего арифметического, моды, медианы, размаха, дисперсии (смещенной и несмещенной). Вывод формулы несмещенной оценки дисперсии. Проверка свойств оценок методом подстановки.	2

2	Математические основы статистического вывода	Интервальное оценивание параметров.	Расчет доверительных интервалов для математического ожидания при известной дисперсии (Z-интервал) и неизвестной дисперсии (t-интервал Стьюдента). Определение необходимого объема выборки для достижения заданной точности оценки. Доверительный интервал для дисперсии через распределение Пирсона.	2
3	Параметрические и непараметрические методы проверки гипотез	Проверка параметрических гипотез.	Формулировка нулевой и альтернативной гипотез. Выбор уровня значимости. Односторонние и двусторонние критические области. Ручной расчет значений Z-статистики, t-статистики Стьюдента и F-статистики Фишера по табличным данным бизнес-кейсов (например, сравнение средних чеков двух магазинов). Принятие решения на основе P-value.	2
4	Параметрические и непараметрические методы проверки гипотез	Непараметрические критерии.	Составление таблицы сопряженности признаков. Расчет теоретических частот и значения критерия Пирсона вручную. Реализация U-критерия Манна-Уитни: ранжирование объединенной выборки, расчет суммы рангов, вычисление эмпирического значения U. Применение критерия знаков.	2
5	Регрессионный анализ и моделирование зависимостей	Метод наименьших квадратов (МНК): алгебра и геометрия.	Вывод системы нормальных уравнений для парной линейной регрессии минимизацией функции потерь RSS (суммы квадратов остатков). Геометрическая интерпретация МНК как ортогональной проекции вектора отклика на пространство столбцов матрицы плана. Расчет коэффициентов регрессии «вручную» для выборки из 3-4 точек.	2
6	Регрессионный анализ и моделирование зависимостей	Свойства оценок МНК и качество модели.	Доказательство несмещенности оценок МНК. Расчет коэффициента детерминации и скорректированного коэффициента детерминации по результатам аналитического примера. Вычисление стандартной ошибки регрессии. Анализ влияния выброса на наклон линии регрессии.	2
7	Вычислительная статистика, регуляризация и снижение размерности	Множественная регрессия и мультиколлинеарность.	Запись матричной формы нормального уравнения. Расчет определителя корреляционной матрицы предикторов для выявления точной мультиколлинеарности. Оценка степени мультиколлинеарности через частные коэффициенты корреляции.	2

8	Вычислительная статистика, регуляризация и снижение размерности	Компромисс смещения и дисперсии (Bias-Variance Tradeoff).	Аналитический разбор графика зависимости ошибки от сложности модели. Математический вывод штрафного слагаемого гребневой регрессии (Ridge/L2). Качественное обсуждение того, почему Lasso (L1) приводит к занулению коэффициентов, а Ridge — только уменьшает их величину. Обсуждение схемы K-Fold кросс-валидации: как именно данные делятся на фолды и как усредняется ошибка.	2
Итого за семестр:				16
Итого:				16

4.4. Содержание самостоятельной работы

Наименование раздела	Вид самостоятельной работы	Содержание самостоятельной работы (перечень дидактических единиц: рассматриваемых подтем, вопросов)	Количество часов
5 семестр			
Математические основы статистического вывода	Математические основы статистического вывода	Проработка лекций, решение задач.	1
Параметрические и непараметрические методы проверки гипотез	Параметрические и непараметрические методы проверки гипотез	статистического вывода Проработка лекций, решение задач.	1
Регрессионный анализ и моделирование зависимостей	Регрессионный анализ и моделирование зависимостей	Проработка лекций, решение задач.	1
Вычислительная статистика, регуляризация и снижение размерности	Вычислительная статистика, регуляризация и снижение размерности	Проработка лекций, решение задач.	1
Итого за семестр:			4
Итого:			4

5. Перечень учебной литературы и учебно-методического обеспечения по дисциплине (модулю)

№ п/п	Библиографическое описание	Ресурс НТБ СамГТУ (ЭБС СамГТУ, IPRbooks и т.д.)
1	Аскеров, П.Ф. Общая и прикладная статистика : учеб. / П. Ф. Аскеров, Р. Н. Пахунова, А. В. Пахунов.- М., Инфра-М, 2014.- 271 с.	Книжный фонд

2	Пугачев, В.С. Теория вероятностей и математическая статистика : Учеб. / В. С. Пугачев .- 2-е изд.,испр.и доп..- М., ФИЗМАТЛИТ, 2002.- 496 с.	Книжный фонд
---	--	--------------

Доступ обучающихся к ЭР НТБ СамГТУ (elib.samgtu.ru) осуществляется посредством электронной информационной образовательной среды университета и сайта НТБ СамГТУ по логину и паролю.

6. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения

При проведении лекционных занятий используется мультимедийное оборудование.

Организовано взаимодействие обучающегося и преподавателя с использованием электронной ин-формационной образовательной среды университета.

№ п/п	Наименование	Производитель	Способ распространения
1	Adobe Reader	Adobe Systems Incorporated (Зарубежный) (Зарубежный)	Свободно распространяемое
2	LibreOffice	The Document Foundation (Зарубежный)	Свободно распространяемое

7. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», профессиональных баз данных, информационно-справочных систем

№ п/п	Наименование	Краткое описание	Режим доступа
-------	--------------	------------------	---------------

8. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Лекционные занятия

- аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук), программное обеспечение: MS Excel, MS Word);
- аудитория, оснащенная учебной мебелью: столы, стулья для обучающихся; стол, стул для преподавателя.

Практические занятия

- аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук), программное обеспечение: MS Excel, MS Word);
- аудитория, оснащенная учебной мебелью: столы, стулья для обучающихся; стол, стул для преподавателя.

Лабораторные занятия

- аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук), программное обеспечение: LibreOffice, Python.
- аудитория, оснащенная учебной мебелью: столы, стулья для обучающихся; стол, стул для преподавателя.

Самостоятельная работа

- компьютерный класс, оснащенный компьютерами с доступом в Интернет и обеспечивающие доступ в электронно-информационную образовательную среду

СамГТУ читальный зал НТБ СамГТУ (ауд. 200 корпус №8; ауд. 125 корпус № 1; ауд. 41, 31, 34, 35
Главный корпус библиотеки; ауд. 83а, 414, 416, 0209 12 корпус; ауд. 401 корпус
№10);

- пакеты ПО общего назначения (MS Excel, MS Word);
- материально-техническое обеспечение НТБ СамГТУ;
- ресурсы УИТ СамГТУ.

9. Методические материалы

Методические рекомендации при работе на лекции

До лекции студент должен просмотреть учебно-методическую и научную литературу по теме лекции с тем, чтобы иметь представление о проблемах, которые будут разбираться в лекции.

Перед началом лекции обучающимся сообщается тема лекции, план, вопросы, подлежащие рассмотрению, доводятся основные литературные источники. Весь учебный материал, сообщаемый преподавателем, должен не просто прослушиваться. Он должен быть активно воспринят, т.е. услышан, осмыслен, понят, зафиксирован на бумаге и закреплён в памяти. Приступая к слушанию нового учебного материала, полезно мысленно установить его связь с ранее изученным. Следя за техникой чтения лекции (акцент на существенном, повышение тона, изменение ритма, пауза и т.п.), необходимо вслед за преподавателем уметь выделять основные категории, законы и определять их содержание, проблемы, предполагать их возможные решения, доказательства и выводы. Осуществляя такую работу, можно значительно облегчить себе понимание учебного материала, его конспектирование и дальнейшее изучение.

Конспектирование лекции позволяет обработать, систематизировать и лучше сохранить полученную информацию с тем, чтобы в будущем можно было восстановить в памяти основные, содержательные моменты. Типичная ошибка, совершаемая обучающимся, дословное конспектирование речи преподавателя. Как правило, при записи «слово в слово» не остается времени на обдумывание, анализ и синтез информации. Отбирая нужную информацию, главные мысли, проблемы, решения и выводы, необходимо сокращать текст, строить его таким образом, чтобы потом можно было легко в нем разобраться. Желательно оставить в рабочих конспектах поля, на которых можно будет делать пометки из рекомендованной литературы, дополняющие материал прослушанной лекции, а также подчеркивающие особую важность тех или иных теоретических положений. С окончанием лекции работа над конспектом не может считаться завершённой. Нужно еще восстановить отдельные места, проверить, все ли понятно, уточнить что-то на консультации и т.п. с тем, чтобы конспект мог быть использован в процессе подготовки к практическим занятиям, зачету, экзамену. Конспект лекции – незаменимый учебный документ, необходимый для самостоятельной работы.

Методические рекомендации при подготовке и работе на практическом занятии

Практические занятия по дисциплине проводятся в целях выработки практических умений и приобретения навыков в решении профессиональных задач.

Рекомендуется следующая схема подготовки к практическому занятию:

1. ознакомление с планом практического занятия, который отражает содержание предложенной темы;
2. проработка конспекта лекции;
3. чтение рекомендованной литературы;
4. подготовка ответов на вопросы плана практического занятия;
5. выполнение тестовых заданий, задач и др.

Подготовка обучающегося к практическому занятию производится по вопросам, разработанным для каждой темы практических занятий и (или) лекций. В процессе подготовки к практическим занятиям, необходимо обратить особое внимание на самостоятельное изучение рекомендованной литературы.

Работа студентов во время практического занятия осуществляется на основе заданий, которые

выдаются обучающимся в начале или во время занятия. На практических занятиях приветствуется активное участие в обсуждении конкретных ситуаций, способность на основе полученных знаний находить наиболее эффективные решения поставленных проблем, уметь находить полезный дополнительный материал по тематике занятий. Обучающимся необходимо обращать внимание на основные понятия, алгоритмы, определять практическую значимость рассматриваемых вопросов. На практических занятиях обучающиеся должны уметь выполнить расчет по заданным параметрам или выработать определенные решения по обозначенной проблеме. Задания могут быть групповые и индивидуальные. В зависимости от сложности предлагаемых заданий, целей занятия, общей подготовки обучающихся преподаватель может подсказать обучающимся алгоритм решения или первое действие, или указать общее направление рассуждений. Полученные результаты обсуждаются с позиций их адекватности или эффективности в рассмотренной ситуации.

Методические рекомендации при работе на лабораторном занятии

Проведение лабораторной работы делится на две условные части: теоретическую и практическую.

Необходимыми структурными элементами занятия являются проведение лабораторной работы, проверка усвоенного материала, включающая обсуждение теоретических основ выполняемой работы.

Перед лабораторной работой, как правило, проводится технико-теоретический инструктаж по использованию необходимого оборудования. Преподаватель корректирует деятельность обучающегося в процессе выполнения работы (при необходимости). После завершения лабораторной работы подводятся итоги, обсуждаются результаты деятельности.

Возможны следующие формы организации лабораторных работ: фронтальная, групповая и индивидуальная. При фронтальной форме выполняется одна и та же работа (при этом возможны различные варианты заданий). При групповой форме работа выполняется группой (командой). При индивидуальной форме обучающимися выполняются индивидуальные работы.

По каждой лабораторной работе имеются методические указания по их выполнению, включающие необходимый теоретический и практический материал, содержащие элементы и последовательную инструкцию по проведению выбранной работы, индивидуальные варианты заданий, требования и форму отчётности по данной работе.

Методические рекомендации по выполнению самостоятельной работы

Организация самостоятельной работы обучающихся ориентируется на активные методы овладения знаниями, развитие творческих способностей, переход от поточного к индивидуализированному обучению с учетом потребностей и возможностей обучающегося.

Самостоятельная работа с учебниками, учебными пособиями, научной, справочной литературой, материалами периодических изданий и Интернета является наиболее эффективным методом получения дополнительных знаний, позволяет значительно активизировать процесс овладения информацией, способствует более глубокому усвоению изучаемого материала. Все новые понятия по изучаемой теме необходимо выучить наизусть и внести в глоссарий, который целесообразно вести с самого начала изучения курса.

Самостоятельная работа реализуется:

- непосредственно в процессе аудиторных занятий;
- на лекциях, практических занятиях;
- в контакте с преподавателем вне рамок расписания;
- на консультациях по учебным вопросам, в ходе творческих контактов, при ликвидации задолженностей, при выполнении индивидуальных заданий и т.д.;
- в библиотеке, дома, на кафедре при выполнении обучающимся учебных и практических задач.

Эффективным средством осуществления обучающимся самостоятельной работы является электронная информационно-образовательная среда университета, которая обеспечивает доступ к учебным планам, рабочим программам дисциплин (модулей), практик, к изданиям электронных библиотечных систем.

10. Фонд оценочных средств по дисциплине (модулю)

Фонд оценочных средств представлен в приложении № 1.

**Фонд оценочных средств
по дисциплине
Б1.О.03.02 «Математическая статистика»**

Код и направление подготовки (специальность)	02.03.03 Математическое обеспечение и администрирование информационных систем
Направленность (профиль)	Анализ данных в бизнес-аналитике
Квалификация	Бакалавр
Форма обучения	Очная
Год начала подготовки	2026
Институт / факультет	Институт инженерно-экономического и гуманитарного образования
Выпускающая кафедра	кафедра "Высшая математика"
Кафедра-разработчик	кафедра "Высшая математика"
Объем дисциплины, ч. / з.е.	108 / 3
Форма контроля (промежуточная аттестация)	Экзамен

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

Наименование категории (группы) компетенций	Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Результаты обучения (знать, уметь, владеть, соотнесенные с индикаторами достижения компетенции)
Общепрофессиональные компетенции			
Теоретические и практические основы профессиональной деятельности	ОПК-1 Способен применять фундаментальные знания, полученные в области математических и (или) естественных наук, и использовать их в профессиональной деятельности	ОПК-1.1 Знать: алгебраические, теоретико-вероятностные и статистические методы для использования в анализе больших данных	Знать алгебраические, теоретико-вероятностные и статистические методы для использования в анализе больших данных.
		ОПК-1.2 Уметь: строить математические модели, отражающие закономерности и взаимосвязи между факторами больших данных	Уметь строить математические модели, отражающие закономерности и взаимосвязи между факторами больших данных.
		ОПК-1.3 Владеть: алгебраическими, теоретико-вероятностными и статистическими понятиями для построения математических моделей	Владеть алгебраическими, теоретико-вероятностными и статистическими понятиями для построения математических моделей.
	ОПК-2 Способен применять современный математический аппарат, связанный с проектированием, разработкой, реализацией и оценкой качества программных продуктов и программных комплексов в различных областях человеческой деятельности	ОПК-2.1 Знать: вычислительные алгоритмы оптимизации и численные методы решения больших линейных систем	Знать вычислительные алгоритмы оптимизации (градиентный спуск, ЕМ-алгоритм) и численные методы решения больших линейных систем (метод наименьших квадратов, сингулярное разложение).

		ОПК-2.2 Уметь: реализовывать различные математические модели (в том числе регрессионные) в виде вычислительных алгоритмов	Уметь реализовывать различные математические модели (в том числе регрессионные) в виде вычислительных алгоритмов с использованием библиотек Python/NumPy/SciPy.
		ОПК-2.3 Владеть: навыками скользящего контроля алгоритма, настроенного на обучающей выборке, и навыками регуляризации алгоритма для предотвращения его переобучения	Владеть навыками скользящего контроля (cross-validation) алгоритма, настроенного на обучающей выборке, и навыками регуляризации алгоритма (L1/L2, Ridge, Lasso) для предотвращения его переобучения.

Матрица соответствия оценочных средств запланированным результатам обучения

Код индикатора достижения компетенции	Результаты обучения	Оценочные средства	Текущий контроль успеваемости	Промежуточная аттестация
Математические основы статистического вывода				
ОПК-1.1 Знать: алгебраические, теоретико-вероятностные и статистические методы для использования в анализе больших данных	Знать алгебраические, теоретико-вероятностные и статистические методы для использования в анализе больших данных.	подготовка к экзамену	Да	Да
ОПК-1.2 Уметь: строить математические модели, отражающие закономерности и взаимосвязи между факторами больших данных	Уметь строить математические модели, отражающие закономерности и взаимосвязи между факторами больших данных.	подготовка к экзамену	Да	Да
ОПК-1.3 Владеть: алгебраическими, теоретико-вероятностными и статистическими понятиями для построения математических моделей	Владеть алгебраическими, теоретико-вероятностными и статистическими понятиями для построения математических моделей.	подготовка к экзамену	Да	Да
ОПК-2.1 Знать: вычислительные алгоритмы оптимизации и численные методы решения больших линейных систем	Знать вычислительные алгоритмы оптимизации (градиентный спуск, EM-алгоритм) и численные методы решения больших линейных систем (метод наименьших квадратов, сингулярное разложение).	подготовка к экзамену	Да	Да

ОПК-2.2 Уметь: реализовывать различные математические модели (в том числе регрессионные) в виде вычислительных алгоритмов	Уметь реализовывать различные математические модели (в том числе регрессионные) в виде вычислительных алгоритмов с использованием библиотек Python/NumPy/SciPy.	подготовка к экзамену	Да	Да
ОПК-2.3 Владеть: навыками скользящего контроля алгоритма, настроенного на обучающей выборке, и навыками регуляризации алгоритма для предотвращения его переобучения	Владеть навыками скользящего контроля (cross-validation) алгоритма, настроенного на обучающей выборке, и навыками регуляризации алгоритма (L1/L2, Ridge, Lasso) для предотвращения его переобучения.	подготовка к экзамену	Да	Да
Параметрические и непараметрические методы проверки гипотез				
ОПК-1.1 Знать: алгебраические, теоретико-вероятностные и статистические методы для использования в анализе больших данных	Знать алгебраические, теоретико-вероятностные и статистические методы для использования в анализе больших данных.	подготовка к экзамену	Да	Да
ОПК-1.2 Уметь: строить математические модели, отражающие закономерности и взаимосвязи между факторами больших данных	Уметь строить математические модели, отражающие закономерности и взаимосвязи между факторами больших данных.	подготовка к экзамену	Да	Да
ОПК-1.3 Владеть: алгебраическими, теоретико-вероятностными и статистическими понятиями для построения математических моделей	Владеть алгебраическими, теоретико-вероятностными и статистическими понятиями для построения математических моделей.	подготовка к экзамену	Да	Да
ОПК-2.1 Знать: вычислительные алгоритмы оптимизации и численные методы решения больших линейных систем	Знать вычислительные алгоритмы оптимизации (градиентный спуск, EM-алгоритм) и численные методы решения больших линейных систем (метод наименьших квадратов, сингулярное разложение).	подготовка к экзамену	Да	Да
ОПК-2.2 Уметь: реализовывать различные математические модели (в том числе регрессионные) в виде вычислительных алгоритмов	Уметь реализовывать различные математические модели (в том числе регрессионные) в виде вычислительных алгоритмов с использованием библиотек Python/NumPy/SciPy.	подготовка к экзамену	Да	Да
ОПК-2.3 Владеть: навыками скользящего контроля алгоритма, настроенного на обучающей выборке, и навыками регуляризации алгоритма для предотвращения его переобучения	Владеть навыками скользящего контроля (cross-validation) алгоритма, настроенного на обучающей выборке, и навыками регуляризации алгоритма (L1/L2, Ridge, Lasso) для предотвращения его переобучения.	подготовка к экзамену	Да	Да
Регрессионный анализ и моделирование зависимостей				

ОПК-1.1 Знать: алгебраические, теоретико-вероятностные и статистические методы для использования в анализе больших данных	Знать алгебраические, теоретико-вероятностные и статистические методы для использования в анализе больших данных.	подготовка к экзамену	Да	Да
ОПК-1.2 Уметь: строить математические модели, отражающие закономерности и взаимосвязи между факторами больших данных	Уметь строить математические модели, отражающие закономерности и взаимосвязи между факторами больших данных.	подготовка к экзамену	Да	Да
ОПК-1.3 Владеть: алгебраическими, теоретико-вероятностными и статистическими понятиями для построения математических моделей	Владеть алгебраическими, теоретико-вероятностными и статистическими понятиями для построения математических моделей.	подготовка к экзамену	Да	Да
ОПК-2.1 Знать: вычислительные алгоритмы оптимизации и численные методы решения больших линейных систем	Знать вычислительные алгоритмы оптимизации (градиентный спуск, ЕМ-алгоритм) и численные методы решения больших линейных систем (метод наименьших квадратов, сингулярное разложение).	подготовка к экзамену	Да	Да
ОПК-2.2 Уметь: реализовывать различные математические модели (в том числе регрессионные) в виде вычислительных алгоритмов	Уметь реализовывать различные математические модели (в том числе регрессионные) в виде вычислительных алгоритмов с использованием библиотек Python/NumPy/SciPy.	подготовка к экзамену	Да	Да
ОПК-2.3 Владеть: навыками скользящего контроля алгоритма, настроенного на обучающей выборке, и навыками регуляризации алгоритма для предотвращения его переобучения	Владеть навыками скользящего контроля (cross-validation) алгоритма, настроенного на обучающей выборке, и навыками регуляризации алгоритма (L1/L2, Ridge, Lasso) для предотвращения его переобучения.	подготовка к экзамену	Да	Да
Вычислительная статистика, регуляризация и снижение размерности				
ОПК-1.1 Знать: алгебраические, теоретико-вероятностные и статистические методы для использования в анализе больших данных	Знать алгебраические, теоретико-вероятностные и статистические методы для использования в анализе больших данных.	подготовка к экзамену	Да	Да
ОПК-1.2 Уметь: строить математические модели, отражающие закономерности и взаимосвязи между факторами больших данных	Уметь строить математические модели, отражающие закономерности и взаимосвязи между факторами больших данных.	подготовка к экзамену	Да	Да

ОПК-1.3 Владеть: алгебраическими, теоретико-вероятностными и статистическими понятиями для построения математических моделей	Владеть алгебраическими, теоретико-вероятностными и статистическими понятиями для построения математических моделей.	подготовка к экзамену	Да	Да
ОПК-2.1 Знать: вычислительные алгоритмы оптимизации и численные методы решения больших линейных систем	Знать вычислительные алгоритмы оптимизации (градиентный спуск, ЕМ-алгоритм) и численные методы решения больших линейных систем (метод наименьших квадратов, сингулярное разложение).	подготовка к экзамену	Да	Да
ОПК-2.2 Уметь: реализовывать различные математические модели (в том числе регрессионные) в виде вычислительных алгоритмов	Уметь реализовывать различные математические модели (в том числе регрессионные) в виде вычислительных алгоритмов с использованием библиотек Python/NumPy/SciPy.	подготовка к экзамену	Да	Да
ОПК-2.3 Владеть: навыками скользящего контроля алгоритма, настроенного на обучающей выборке, и навыками регуляризации алгоритма для предотвращения его переобучения	Владеть навыками скользящего контроля (cross-validation) алгоритма, настроенного на обучающей выборке, и навыками регуляризации алгоритма (L1/L2, Ridge, Lasso) для предотвращения его переобучения.	подготовка к экзамену	Да	Да

Контрольные задания и оценочные материалы для текущего контроля успеваемости по дисциплине «Математическая статистика», сгруппированные по формируемым компетенциям (ОПК-1 и ОПК-2).

1. Оценочные средства для контроля знаний (ОПК-1.1, ОПК-2.1)

Форма: Тестирование в системе дистанционного обучения (LMS) по итогам каждого лекционного модуля.

- **Пример задания (выбор одного варианта):** при каком условии оценка метода наименьших квадратов (МНК) является наилучшей линейной несмещенной оценкой (BLUE) согласно теореме Гаусса-Маркова?
 - о А) Отсутствие автокорреляции и гомоскедастичность ошибок.
 - о Б) Нормальное распределение зависимой переменной.
 - о В) Отсутствие мультиколлинеарности независимых переменных.
 - о Г) Большой объем выборки ($n > 30$). *(Правильный ответ: А)*
- **Пример задания (на соответствие):** Соотнесите вычислительный алгоритм и задачу, которую он решает:
 - 1. QR-разложение
 - 2. Метод градиентного спуска
 - 3. Сингулярное разложение (SVD) А) Итеративная минимизация функции потерь в больших линейных системах Б) Поиск устойчивого решения системы нормальных уравнений В) Понижение размерности и сжатие данных *(Ответ: 1-Б, 2-А, 3-В)*

2. Оценочные средства для контроля умений (ОПК-1.2, ОПК-2.2)

Форма: Решение практических задач (case-study) на семинарах или в домашних работах.

- **Пример задания (бизнес-кейс):** *Условие:* Аналитiku предоставлены данные о продажах мороженого (Y , тыс. руб.) за 50 дней, средней дневной температуре (X_1 , °C) и затратах на рекламу (X_2 , тыс. руб.). *Задание:*
 - 1. Построить матрицу парных корреляций. Сделать вывод о наличии мультиколлинеарности.
 - 2. Записать спецификацию множественной линейной регрессионной модели.
 - 3. Рассчитать коэффициенты модели аналитически.
 - 4. Интерпретировать экономический смысл коэффициента при переменной X_1 .

5. Проверить статистическую значимость модели в целом (F-критерий Фишера) и каждого коэффициента в отдельности (t-критерий Стьюдента) на уровне значимости 0.05.

3. Оценочные средства для контроля навыков и опыта деятельности (ОПК-1.3, ОПК-2.3)

Форма: Лабораторные работы с обязательной защитой (демонстрация кода и результатов).

Лабораторная работа № 4: «Реализация и регуляризация регрессионных моделей»

- **Задача:** написать с нуля на Python (библиотека NumPy) функцию, реализующую линейную регрессию с L2-регуляризацией (Ridge).
- **Требования к коду:**
 - Функция должна принимать обучающую матрицу X , вектор y и коэффициент регуляризации α .
 - Реализовать вычисление весов через аналитическое решение с учетом регуляризации.
- **Требования к анализу (опыт деятельности):**
 1. Разделить предоставленный датасет на обучающую (70%) и тестовую (30%) выборки.
 2. Настроить гиперпараметр α с помощью **скользящего контроля (кросс-валидации)** (K-Fold, $K=5$). Построить график зависимости ошибки кросс-валидации от значения α .
 3. Обучить финальную модель на всем обучающем наборе с оптимальным α .
 4. Рассчитать метрики RMSE и R^2 на тестовой выборке.
 5. **Задание на проверку навыка предотвращения переобучения:** Продемонстрировать на графике (кривые обучения), что при $\alpha = 0$ модель имеет разрыв между ошибкой на обучении и ошибкой на кросс-валидации (переобучение), а при оптимальном α этот разрыв минимален.

Лабораторная работа № 6: «Сравнение алгоритмов оптимизации»

- **Задача:** Реализовать градиентный спуск (с подбором learning rate) и сравнить его с аналитическим решением МНК на синтетическом датасете из 100 000 наблюдений. Оценить время работы и точность (сходимость) обоих методов.

4. Материалы для контроля самостоятельной работы

Форма: Контрольные вопросы для самопроверки и подготовки к коллоквиуму (внеаудиторная КСР):

1. В чем разница между состоятельностью и несмещенностью статистической оценки? Приведите пример состоятельной, но смещенной оценки.
2. Почему при наличии гетероскедастичности стандартные ошибки МНК становятся некорректными? Какой метод используется для их исправления (White's robust standard errors)?
3. Объясните геометрический смысл регуляризации L1 (Lasso) и L2 (Ridge) в пространстве коэффициентов регрессии. Почему L1 способствует отбору признаков (делает веса точно равными нулю)?
4. Как центральная предельная теорема обосновывает использование параметрических критериев (t-теста) при неизвестном распределении генеральной совокупности?

Критерии оценивания текущего контроля:

- **Тесты (LMS):** Автоматическая проверка. Зачет ставится при правильном ответе на 80% и более вопросов.
- **Практические задачи:** Оцениваются полнота математического вывода, корректность интерпретации статистических показателей (p-value, доверительные интервалы) и отсутствие логических ошибок.
- **Лабораторные работы:** Оцениваются по чек-листу:
 - Работоспособность и читаемость кода (30%).
 - Корректность применения методов регуляризации и кросс-валидации (40%).
 - Глубина анализа результатов и бизнес-интерпретация (30%).

Формы промежуточной аттестации

Вопросы к экзамену (семестр 5)

1. Понятие случайной величины. Законы распределения: нормальное, распределение Стьюдента, хи-квадрат, Фишера. Их роль в статистическом выводе.
2. Выборочный метод. Генеральная и выборочная совокупности. Репрезентативность выборки. Вариационный ряд.
3. Точечные оценки параметров распределения. Свойства оценок: несмещенность, состоятельность, эффективность.
4. Выборочное среднее и выборочная дисперсия как оценки математического ожидания и дисперсии.

5. Интервальное оценивание. Построение доверительного интервала для математического ожидания при известной и неизвестной дисперсии.
6. Доверительный интервал для дисперсии нормального распределения.
7. Статистическая гипотеза. Нулевая и альтернативная гипотезы. Ошибки первого и второго рода. Уровень значимости и мощность критерия.
8. Проверка гипотез о числовых значениях параметров (t-критерий Стьюдента для одной и двух выборок).
9. Проверка гипотез о равенстве дисперсий (F-критерий Фишера).
10. Критерии согласия: критерий хи-квадрат Пирсона и критерий Колмогорова-Смирнова. Проверка нормальности распределения данных.
11. Однофакторный дисперсионный анализ (ANOVA). Сравнение средних нескольких групп.
12. Непараметрические критерии: критерий знаков, U-критерий Манна-Уитни (применение в бизнес-аналитике при нарушении условий нормальности).
13. Корреляционный анализ. Ковариация и коэффициент корреляции Пирсона.
14. Проверка значимости коэффициента корреляции. 14. Ранговая корреляция (коэффициенты Спирмена и Кендалла).
15. Парная линейная регрессия. Метод наименьших квадратов (МНК). Вывод системы нормальных уравнений.
16. 1 Анализ качества парной регрессии: коэффициент детерминации,
17. . Предпосылки классической линейной модели (теорема Гаусса-Маркова). Гомоскедастичность и гетероскедастичность.
18. Мультиколлинеарность. Причины возникновения, методы обнаружения (VIF) и способы устранения.
19. Множественная линейная регрессия. Интерпретация коэффициентов регрессии в бизнес-контексте.
20. Нелинейная регрессия. Полиномиальная регрессия и использование признаков в полиномиальных степенях.
21. Матричная форма записи МНК. Решение системы нормальных уравнений. QR-разложение.
22. Градиентные методы оптимизации функции потерь (Batch, Stochastic, Mini-batch GD). Скорость сходимости.
23. Переобучение (overfitting) и недообучение (underfitting). Компромисс между смещением и дисперсией (Bias-Variance Tradeoff).
24. Регуляризация линейных моделей. L2-регуляризация (Ridge regression). Геометрическая интерпретация и формула оценки весов.
25. L1-регуляризация (Lasso regression). Свойство отбора признаков (обнуление весов).
26. Скользящий контроль (кросс-валидация). Метод K-Fold. Оценка обобщающей способности модели.
27. Метрики качества регрессионных моделей: MSE, RMSE, MAE, MAPE. Области их применимости в бизнес-задачах.

28. Логистическая регрессия. Логистическая функция (сигмоида). Оценка параметров через метод максимального правдоподобия.
29. Понятие статистической значимости признаков в моделях машинного обучения (p -value коэффициентов, отбор признаков).
30. Bootstrap (бутстреп-анализ) как метод оценки статистической устойчивости параметров модели.

Экзамен проводится в устной и письменной форме. Экзаменационный билет содержит 3 вопроса из различных разделов дисциплины, из которых 2 – теоретических и 1 – практический.

Практические задания к экзамену составлены аналогично заданиям из контрольных работ по всем темам изученных разделов дисциплины.

Пример формирования билета:

Федеральное государственное
бюджетное образовательное
учреждение высшего
образования Самарский
государственный технический
университет



ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 1 ПО ДИСЦИПЛИНЕ «Математическая статистика»

ИИЭиГО

Кафедра высшей математики Курс 3 (семестр 5)

Направление: 02.03.03 Математическое
обеспечение и администрирование
информационных систем

Профиль подготовки: Анализ данных в бизнес-
аналитике

СОСТАВ БИЛЕТА

1. Предпосылки классической линейной модели (теорема Гаусса-Маркова). Последствия гетероскедастичности для оценок МНК. Тест Уайта.
2. L2-регуляризация (Ridge regression). Запишите целевую функцию метода и аналитическое решение для вектора весов. Объясните, как параметр регуляризации влияет на величину коэффициентов и почему Ridge, в отличие от Lasso, не обнуляет веса.
3. Вам предоставлен датасет sales_data.csv, содержащий 500 наблюдений. Столбцы: Sales (объем продаж, целевая переменная), TV (бюджет на ТВ-рекламу), Radio (бюджет на радио), Newspaper (бюджет на газеты). Все бюджеты указаны в тыс. долларов.

Задание:

1. Загрузите данные и постройте матрицу корреляций. Существует ли проблема мультиколлинеарности между факторами? Обоснуйте ответ.

2. Реализуйте модель множественной линейной регрессии для прогнозирования Sales (используйте scikit-learn). Запишите уравнение полученной модели.
3. Проведите **скользящий контроль (K-Fold кросс-валидацию с K=5)** для оценки качества модели. Вычислите среднее значение метрики RMSE.
4. Настройте гиперпараметр регуляризации alpha для Ridge-регрессии с помощью GridSearchCV. Какой alpha оказался оптимальным?
5. Обучите финальную модель Ridge с оптимальным alpha на всем обучающем наборе. Сравните значения коэффициентов с моделью из п. 2. Сделайте вывод о том, как регуляризация повлияла на веса признаков Radio и Newspaper.
6. Визуализируйте график зависимости RMSE кросс-валидации от значения alpha (график валидации). Покажите на нем точку переобучения (если она есть) и точку оптимального alpha.

СОСТАВИЛ к.ф.-м.н., доцент кафедры ВМ “УТВЕРЖДАЮ” зав. кафедрой ВМ
ШУВАЛОВА Т.В. ЮСУПОВА О.В.

Дата:

Дата:

Тесты по компетенциям

Типовые задания для промежуточной аттестации по дисциплине

Дисциплина Б1.О.03.02 «Математическая статистика»

Направление подготовки 02.03.03 Математическое обеспечение и администрирование информационных систем

Профиль Анализ данных в бизнес-аналитике

№ задания	Содержание задания	Ключ к ответу	Наименование темы РПД	Семестр
ОПК-1 Способен применять фундаментальные знания, полученные в области математических и (или) естественных наук, и использовать их в профессиональной деятельности				
1	Какое распределение используется для построения доверительного интервала для математического ожидания при неизвестной дисперсии выборки? А) Нормальное	Б	Точечные и интервальные оценки параметров распределения	5

	Б) Стьюдента В) Хи-квадрат Г) Фишера			
2	Что характеризует p-value при проверке статистических гипотез? А) Вероятность того, что нулевая гипотеза верна Б) Вероятность получить наблюдаемые или более экстремальные данные при условии, что нулевая гипотеза верна В) Мощность статистического критерия Г) Уровень значимости	Б	Проверка статистических гипотез	5
3	Если коэффициент корреляции Пирсона $r = -0.92$, какой вывод можно сделать о связи между факторами? А) Связь отсутствует Б) Слабая положительная связь В) Сильная обратная (отрицательная) линейная связь Г) Причинно-следственная связь	В	Корреляционный анализ	5
4	Какое условие классической линейной модели (МНК) нарушается, если дисперсия остатков модели не постоянна, а зависит от значений факторов? А) Отсутствие автокорреляции Б) Гомоскедастичность В) Линейность параметров Г) Отсутствие мультиколлинеарности	Б	Предпосылки метода наименьших квадратов (Гаусса-Маркова)	5

5	Что является основным негативным последствием мультиколлинеарности в модели множественной регрессии? А) Смещение оценок коэффициентов Б) Занижение коэффициента детерминации В) Завышение стандартных ошибок коэффициентов, делающее их статистически незначимыми Г) Невозможность вычислить обратную матрицу	В	Мультиколлинеарность и способы её устранения	5
6	Какая метрика качества регрессии (функция потерь) более устойчива к выбросам в данных по сравнению с MSE (среднеквадратичной ошибкой)? А) R^2 Б) MAE (средняя абсолютная ошибка) В) Логарифмическая функция потерь Г) MSE	Б	Метрики качества регрессионных моделей	5
7	Какое свойство точечной оценки гарантирует, что при увеличении объема выборки до бесконечности оценка сходится по вероятности к истинному значению параметра? А) Несмещенность, Б) Эффективность, В) Состоятельность, Г) Достаточность	В	Точечные оценки параметров распределения	5
8	Для чего в бизнес-аналитике применяется	А	Проверка статистических	5

	F-критерий Фишера в контексте регрессионного анализа? А) Для проверки значимости всей модели в целом, Б) Для проверки нормальности распределения остатков, В) Для обнаружения гетероскедастичности, Г) Для отбора признаков		гипотез (Общая значимость регрессии)	
9	При применении L1-регуляризации (Lasso) к модели линейной регрессии: А) Все коэффициенты остаются ненулевыми, Б) Некоторые коэффициенты могут стать точно равными нулю (происходит отбор признаков), В) Сумма квадратов коэффициентов добавляется к функции потерь, Г) Оценки коэффициентов всегда становятся смещенными в сторону увеличения	Б	Регуляризация линейных моделей (L1 и L2)	5
10	Какое распределение вероятностей лежит в основе проверки статистических гипотез о дисперсии (например, при расчете доверительного интервала для σ^2) А) Нормальное, Б) Стьюдента, В) Хи-квадрат (χ^2), Г) Равномерное	В	Точечные и интервальные оценки параметров распределения	5
ОПК-2 Способен применять современный математический аппарат, связанный с проектированием, разработкой, реализацией и оценкой качества программных продуктов и программных комплексов в различных областях человеческой деятельности				

1	Какой вычислительный метод решения системы нормальных уравнений $(X^T X)\beta = X^T Y$ является наиболее численно устойчивым при работе с большими данными и мультиколлинеарностью, в отличие от прямого вычисления обратной матрицы? А) Метод Гаусса, Б) Сингулярное разложение (SVD), В) QR-разложение матрицы признаков, Г) Метод Якоби	В	Вычислительные методы линейной алгебры в статистике	5
2	В чем заключается основная вычислительная сложность (узкое место) при обучении модели линейной регрессии на датасете с 1 000 000 наблюдений и 500 признаками методом нормальных уравнений? А) Вычисление транспонированной матрицы $X^T X$, Б) Обращение матрицы признаков размерности 501×501 , В) Перемножение вектора на вектор, Г) Сортировка данных по признаку	Б	Вычислительные методы линейной алгебры в статистике	5
3	Какое фундаментальное свойство алгоритма проверяет процедура K-Fold кросс-валидации (скользящего контроля)? А) Скорость сходимости градиентного спуска, Б) Обобщающую способность модели и риск переобучения (overfitting), В) Наличие пропусков в данных, Г)	Б	Кросс-валидация и оценка качества алгоритмов	5

	Линейную зависимость признаков			
4	Какой метод регуляризации (из перечисленных) вычислительно используется не только для снижения дисперсии модели, но и как алгоритм отбора признаков (feature selection), обнуляя веса наименее важных переменных? А) Ridge (L2), Б) Lasso (L1), В) Elastic Net, Г) Early Stopping	Б	Регуляризация алгоритмов (L1, L2, Elastic Net)	5
5	При реализации алгоритма линейной регрессии методом градиентного спуска на невыпуклой функции потерь или при наличии зашумленных данных, какой вариант обновления весов обеспечивает более стабильную сходимость за счет сглаживания шума? А) Batch Gradient Descent, Б) Stochastic Gradient Descent (SGD), В) Mini-batch Gradient Descent, Г) Метод Ньютона	В	Вычислительные алгоритмы оптимизации	5
6	Какое явление описывает ситуация, когда ошибка модели на обучающей выборке близка к нулю, а на тестовой (валидационной) выборке — очень высока? С точки зрения реализации, это означает, что алгоритм «запомнил» шум данных вместо	Б	Компромисс между смещением и дисперсией (Bias-Variance Tradeoff)	5

	закономерностей. А) Недообучение (Underfitting), Б) Переобучение (Overfitting), В) Гетероскедастичность, Г) Смещение оценок			
7	Какая техника регуляризации алгоритма (помимо L1/L2 на коэффициенты) применяется в обучении ансамблей (например, градиентного бустинга) для предотвращения переобучения путем ограничения роста деревьев? А) Увеличение скорости обучения (learning rate), Б) Early Stopping (ранняя остановка по метрике валидации), В) Увеличение глубины деревьев, Г) Отбор признаков вручную	Б	Регуляризация алгоритмов и предотвращение переобучения	5
8	При программной реализации функции потерь MSE (Mean Squared Error) для задачи регрессии, какая численная проблема может возникнуть при возведении больших ошибок в квадрат и как ее обычно решают в коде для предотвращения переполнения? А) Проблема исчезает при нормализации целевой переменной y, Б) Проблема линейной алгебры, В) Проблема решается увеличением batch size, Г) Проблема решается использованием L1-ошибки (MAE) или	А, Г	Реализация вычислительных алгоритмов и функции потерь	5

	обрезкой градиентов (clipping)			
9	Какой векторизованный подход на Python (NumPy/Pandas) является наиболее эффективным с точки зрения использования оперативной памяти и скорости процессора для вычисления прогнозов $\hat{Y} = X\beta$ на батче из 10 000 объектов? А) Цикл for по строкам матрицы X , Б) Использование списковых включений (list comprehension), В) Матричное умножение через np.dot(X, beta) или оператор @, Г) Применение функции apply() к датафрейму	В	Программная реализация математических моделей	5
10	Какая метрика качества (loss function) является дифференцируемой и наиболее предпочтительной для использования в качестве функции потерь при настройке весов модели методом градиентного спуска? А) MAE (Mean Absolute Error), Б) MSE (Mean Squared Error), В) Huber Loss (при больших ошибках ведет себя как MSE, при малых — как MAE), Г) Точность (Accuracy)	Б, В	Оценка качества программных продуктов (моделей)	5

Методические материалы, определяющие процедуры оценивания для дисциплины «Математическая статистика».

Оценивание компетенций ОПК-1 и ОПК-2 проводится непрерывно в рамках балльно-рейтинговой системы (максимум 100 баллов).

1. Текущий контроль (максимум 60 баллов, порог допуска к экзамену — 40 баллов):

- **ОПК-1.1, ОПК-1.3 (Знания):** Проверяются тестами в LMS после каждого лекционного модуля (10 баллов).
- **ОПК-1.2, ОПК-2.2 (Умения):** Проверяются на практических занятиях решением задач на построение моделей (10 баллов).
- **ОПК-2.1, ОПК-2.2, ОПК-2.3 (Владение/Навыки):** Проверяются на защите лабораторных работ. Каждая из 4 двухчасовых работ оценивается по 10 баллов. Оценка учитывает корректность кода, применение регуляризации (L1/L2) и использование кросс-валидации (скользящего контроля) для оценки переобучения (40 баллов).

2. Промежуточная аттестация (Экзамен, максимум 40 баллов): Проводится в комбинированной форме (компьютерное тестирование + решение практической задачи).

- **ОПК-1.1, ОПК-2.1:** Тестовые вопросы на знание определений, формул и алгоритмов (20 баллов).
- **ОПК-1.2, ОПК-2.2:** Практическое задание. Студенту предоставляется датасет. Требуется построить и обучить регрессионную модель, подобрать гиперпараметры регуляризации, оценить качество через кросс-валидацию и интерпретировать результаты (20 баллов).

Шкала перевода итогового рейтинга в оценку:

- 86–100 баллов — «Отлично»
- 71–85 баллов — «Хорошо»
- 51–70 баллов — «Удовлетворительно»
- 0–50 баллов — «Неудовлетворительно»