

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Механико-математический факультет

УТВЕРЖДЕНО:

Декан

Л.В. Гензе

Оценочные материалы по дисциплине

Прикладной статистический анализ

по направлению подготовки

01.04.01 Математика

Направленность (профиль) подготовки:
Моделирование и цифровые двойники

Форма обучения

Очная

Квалификация

Магистр

Год приема

2025

СОГЛАСОВАНО:

Руководитель ОП

Е.И. Гурина

Председатель УМК

Е.А. Тарасов

Томск – 2025

1. Компетенции и индикаторы их достижения, проверяемые данными оценочными материалами

Целью освоения дисциплины является формирование следующих компетенций:

ОПК-1 Способен формулировать и решать актуальные и значимые проблемы математики.

ОПК-2 Способен строить и анализировать математические модели в современном естествознании, технике, экономике и управлении.

ПК-1 Способен разрабатывать и внедрять цифровые двойники, используя современные технологии, методы и инструменты, с учетом технических требований заказчика и специфики моделируемых объектов и процессов..

ПК-2 Способен проводить тестирование, валидацию и анализ данных цифровых двойников для обеспечения их корректной работы, оптимизации процессов и принятия решений..

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК 1.1 Формулирует поставленную задачу, пользуется языком предметной области, обоснованно выбирает метод решения задачи.

ИОПК 1.2 Анализирует актуальные и значимые проблемы математики и существующие подходы к их решению.

ИОПК 2.1 Анализирует, выбирает и обосновывает математические модели для решения задач в области современного естествознания, техники, экономики и управления.

ИОПК 2.2 Разрабатывает новые и/или адаптирует/совершенствует математические модели для задач современного естествознания, техники, экономики и управления под руководством более квалифицированного работника.

ИПК 1.1 Анализирует и выбирает современные технологии, методы и инструменты для проектирования и разработки цифровых двойников с учетом специфики решаемых задач.

ИПК 1.3 Разрабатывает математические модели и алгоритмы для создания математической основы цифровых двойников изделий и технических систем.

ИПК 2.2 Анализирует и интерпретирует данные, полученные от цифровых двойников, для принятия предиктивных решений и оптимизации процессов.

2. Оценочные материалы текущего контроля и критерии оценивания

Элементы текущего контроля:

– практические задания.

Примерные темы практических занятий с примерами заданий:

- *Практическая работа №1 «Точечные оценки, свойства и доверительные интервалы»* (ИОПК 1.1, ИОПК 1.2, ИПК 1.1, ИПК 2.2)
- *Практическая работа №2 «Линейная регрессия (парная и множественная)»* (ИОПК 1.2, ИОПК 2.1, ИОПК 2.2, ИПК 1.1, ИПК 1.3, ИПК 2.2)
- *Практическая работа №3 «Регуляризация и логистическая регрессия»* (ИОПК 1.2, ИОПК 2.1, ИОПК 2.2, ИПК 1.1, ИПК 1.3, ИПК 2.2)
- *Практическая работа №4 «Задача о разладке»* (ИОПК 1.1, ИОПК 1.2, ИПК 1.1, ИПК 2.2)

Практическая работа №1 «Точечные оценки, свойства и доверительные интервалы»

Работа выполняется на языке Python. Вам необходимо провести полный статистический анализ данных, включая точечное оценивание и построение доверительных интервалов.

Примеры заданий:

- Сгенерируйте в Python выборку объемом 100 наблюдений, смоделированную из показательного распределения с параметром $\lambda = 0.1$ или найдите реальные данные.
- Для полученной выборки решите следующие задачи:
 - 1) Показательное распределение: Предположив, что данные были сгенерированы из показательного распределения с параметром λ , найдите точечную оценку этого параметра, используя метод моментов.
 - 2) Нормальное распределение: Предположив, что данные были сгенерированы из нормального распределения $N(\mu, \sigma^2)$, найдите точечные оценки для математического ожидания μ и дисперсии σ^2 , используя метод моментов.
- Используя возможности Python для статистических расчетов, постройте доверительные интервалы(ДИ):
 - 1) ДИ для математического ожидания нормального распределения (с надежностью 95%), предполагая, что: Дисперсия генеральной совокупности известна и равна выборочной дисперсии; Дисперсия генеральной совокупности неизвестна.
 - 2) ДИ для параметра λ показательного распределения (с надежностью 95%)
 - 3) ДИ для дисперсии нормального распределения (с надежностью 90%).
- Анализ и интерпретация результатов
 - 1) Сравните полученную оценку параметра λ с его истинным значением. Сделайте вывод о точности оценки. Обоснуйте, почему выбранные вами данные могут описываться показательным распределением.
 - 2) Анализ адекватности моделей: Насколько обоснованным является предположение о нормальности распределения для ваших данных?
 - 3) Интерпретация доверительных интервалов: Дайте содержательную интерпретацию построенным доверительным интервалам. Сравните ширину доверительных интервалов для случаев известной и неизвестной дисперсии. Проанализируйте, как объем выборки влияет на точность оценок

Практическая работа №2: «Линейная регрессия (парная и множественная)»

Освоить методы построения и анализа линейных регрессионных моделей.

Примеры заданий:

Используя Python и библиотеки pandas, statsmodels и scikit-learn, выполните следующее:

Найти любой датасет с количественными переменными. Выбрать и исследовать зависимости различных показателей, для этого

- Постройте модель парной линейной регрессии.
- Дайте содержательную интерпретацию.
- Постройте модель множественной регрессии.
- Проведите сравнительный анализ моделей.
- Проведите диагностику и верификацию лучшей модели.

Практическая работа №3: «Регуляризация и логистическая регрессия»

Освоить методы работы с мультиколлинеарностью с помощью регуляризации и изучить основы бинарной классификации с использованием логистической регрессии.

Примеры заданий:

- Сгенерируйте в Python синтетический набор мультиколлинеарных данных.
- Диагностика мультиколлинеарности. Постройте обычную линейную регрессию и проанализируйте устойчивость оценок.
- Реализуйте Ridge-регрессию для сгенерированных данных.

- Сравните обычную линейную регрессию и Ridge-регрессию.
- Постройте логистическую регрессию для предсказания вероятности.

Практическая работа №4: «Задача о разладке»

Освоить методы статистического контроля качества и обнаружения разладки в стационарных процессах.

Примеры заданий:

Используя Python, сгенерируйте данные диаметра валиков на производстве:

- Первые 20 точек: стабильный процесс $\sim N(50.0, 0.1)$
- С 21 по 30 точку: появление разладки $\sim N(50.8, 0.1)$

Задачи:

- Рассчитайте выборочное среднее и стандартное отклонение по первым 20 точкам ("фаза I").
- Постройте контрольную карту для индивидуальных значений (I-chart) с границами $\mu \pm 3\sigma$.
- Определите момент выхода процесса из-под контроля.
- Рассчитайте кумулятивные суммы (CUSUM) для обнаружения малых смещений и проведите анализ эффективности алгоритма CUSUM.

Характерными показателями развития самостоятельности у студента в результате освоения дисциплины являются: теоретическое осмысление изучаемого материала, накопление необходимых умений и навыков, интерес к процессу создания продукта собственной самостоятельной деятельности, умение провести презентацию созданного продукта, умение отстаивать собственную точку зрения или предложенный вариант решения проблемы, рефлексия своей деятельности и результата.

Критерии оценивания:

Текущий контроль по дисциплине проводится путем контроля посещаемости, оценки отчетов по практическим занятиям.

Подготовка к практическим занятиям предполагает самостоятельную работу студентов по проведению расчетов, анализу, обработке данных, оформлению отчетов. Балльная оценка текущего контроля успеваемости студента по данной дисциплине составляет максимум 56 баллов.

Таблица 1

№ п/п	Вид контроля	Количество	Количество баллов за 1 ед. контроля	Сумма баллов
1.	Посещение лекций	8	0-1	0-8
2.	Выполнение практических заданий	4	0-12	0-48
	Итого			0-56

Основным критерием балльной оценки текущего контроля успеваемости является оценка качества выполнения практического задания (содержание ответа, полнота ответа, владение профессиональным языком).

Индикаторы балльной оценки практических заданий:

- 10-12 баллов – ответ не содержит ошибочных расчетов, элементов и утверждений, максимально полно раскрывает суть каждого вопроса, составлен профессиональным языком, содержит выводы;

- 7-9 баллов – в ответе допущены непринципиальные ошибки и неточности в расчетах, ответ содержит упущения, составлен профессиональным языком, содержит выводы;
- 4-6 баллов – ответ содержит несколько ошибок в расчетах, упущения, содержание ответов не полное; составлен профессиональным языком, в выводах допущены неточности;
- 0-3 баллов – ответ содержит многочисленные ошибки в расчетах, упущения, содержание ответов не полное; выводы отсутствуют.

3. Оценочные материалы итогового контроля (промежуточной аттестации) и критерии оценивания

Итоговый контроль проводится по результатам текущего контроля с выставлением дифференцированного зачета.

Соответствие 56-балльной шкалы оценок 4-альтернативной шкале оценок:

- 0-25 баллов – «неудовлетворительно»,
- 26-40 баллов – «удовлетворительно»,
- 41-50 баллов – «хорошо»,
- 51-56 баллов – «отлично».

4. Оценочные материалы для проверки остаточных знаний (сформированности компетенций)

Теоретические вопросы:

1. Точечное оценивание параметров распределений. (ИОПК-1.1, ИОПК-2.1)
 Ответ должен содержать: определение точечной оценки; методы нахождения оценок (метод моментов, максимального правдоподобия); свойства «хороших» оценок (несмещенность, эффективность, состоятельность).
2. Доверительные интервалы для параметров распределений. (ИОПК-2.1, ИПК-2.2)
 Ответ должен содержать: понятие доверительного интервала и доверительной вероятности; построение доверительных интервалов для математического ожидания и дисперсии нормального распределения; интерпретацию полученных интервалов.
3. Парная и множественная линейная регрессия. (ИОПК-2.2, ИПК-1.3)
 Ответ должен содержать: постановку задачи регрессионного анализа; метод наименьших квадратов для оценки параметров; проверку значимости коэффициентов и модели в целом; коэффициенты детерминации и их интерпретацию.
4. Диагностика регрессионных моделей. (ИПК-2.2)
 Ответ должен содержать: проверку условий Гаусса-Маркова (гомоскедастичность, отсутствие автокорреляции, нормальность остатков); анализ выбросов и влиятельных наблюдений; последствия нарушения предпосылок.
5. Мультиколлинеарность и методы регуляризации. (ИПК-1.1, ИПК-1.3)
 Ответ должен содержать: понятие и причины мультиколлинеарности; методы диагностики (VIF, корреляционная матрица); суть ридж-регрессии (L2-регуляризация) и отличия от метода наименьших квадратов.
6. Логистическая регрессия. (ИОПК-2.2, ИПК-1.3)
 Ответ должен содержать: постановку задачи бинарной классификации; вид и свойства логистической функции; интерпретацию коэффициентов через

отношения шансов (odds ratio); оценку качества классификации (матрица ошибок, ROC-кривая).

7. Задача о разладке и статистический контроль процессов. (ИОПК-2.1, ИПК-2.2)

Ответ должен содержать: постановку задачи обнаружения разладки; кумулятивные суммы (CUSUM) для обнаружения малых смещений; интерпретацию карт контроля качества.

8. Применение Python для статистического анализа данных. (ИОПК-1.1, ИПК-1.1)

Ответ должен содержать: обзор библиотек (pandas, numpy, scipy, statsmodels, scikit-learn) и их функционала для решения задач прикладной статистики; примеры использования для построения регрессионных моделей и проверки гипотез.

Информация о разработчиках

Губин Владимир Николаевич, к. ф.-м. н., доцент кафедры математического анализа и теории функций,

Лещинская Мария Анатольевна, ст. преп. кафедры математического анализа и теории функций,

Пчелинцев Евгений Анатольевич, д.ф.-м.н., доцент, зав. каф. кафедрой математического анализа и теории функций.

Курс «**Прикладной статистический анализ**» разработан при поддержке Регионального научно-образовательного математического центра Томского государственного университета ([НОМЦ ТГУ](#)).