

МИНОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Институт прикладной математики и компьютерных наук

УТВЕРЖДАЮ

Директор института прикладной
математики и компьютерных наук

А.В. Замятин

« 11 » _____ 2021 г.



Фонд оценочных средств по дисциплине

Статистические методы машинного обучения

Направление подготовки

01.03.02 Прикладная математика и информатика

Профиль: **Прикладная математика и информатика**

ОС составила:

канд. физ.-мат. наук, доцент,
доцент кафедры теории вероятностей
и математической статистики



Т.В. Кабанова

Рецензент:

канд. физ.-мат. наук,
доцент кафедры теории вероятностей
и математической статистики



Е.Ю. Лисовская

Фонд оценочных средств одобрен на заседании учебно-методической комиссии
института прикладной математики и компьютерных наук (УМК ИПМКН).

Протокол от 17 июня 2021 г. № 05

Председатель УМК ИПМКН,
д-р техн. наук, профессор



С.П. Сущенко

Оценочные средства (ОС) являются элементом системы оценивания сформированности компетенций у обучающихся в целом или на определенном этапе ее формирования.

ОС разрабатывается в соответствии с рабочей программой (РП) дисциплины.

1. Компетенции и результаты обучения, формируемые в результате освоения дисциплины

Компетенция	Индикатор компетенции	Код и наименование результатов обучения (планируемые результаты обучения, характеризующие этапы формирования компетенций)	Критерии оценивания результатов обучения			
			Отлично	Хорошо	Удовлетворительно	Неудовлетворительно
ОПК-3. Способен применять и модифицировать математические модели для решения задач в области профессиональной деятельности	<p>ИОПК-3.1. Демонстрирует навыки применения современного математического аппарата для построения адекватных математических моделей реальных процессов, объектов и систем в своей предметной области.</p> <p>ИОПК-3.2. Демонстрирует умение собирать и обрабатывать статистические, экспериментальные, теоретические и т.п. данные для построения математических моделей, расчетов и конкретных практических выводов.</p>	<p>ОР-3.1. Обучающийся будет знать методы подбора и обработки информации относительно выбранной темы исследования, методы построения и анализа математических моделей процессов, относящихся к различным предметным областям.</p> <p>ОР-3.2. Обучающийся будет уметь находить в специальной литературе необходимую информацию по соответствующей проблеме, выбрать метод исследования соответствующий поставленной цели с учетом априорных знаний об исследуемом объекте.</p> <p>ОР-3.3. Обучающийся будет владеть навыками решения практических задач и анализа полученных результатов, навыками формализации и построения алгоритма решения исследовательских задач.</p>	Сформированные систематические знания умения и навыки.	Сформированные, но содержащие отдельные пробелы знания, и в целом успешное, но сопровождающееся отдельными ошибками умения и навыки.	Общие, но не структурированные знания и, в целом успешные, но не систематически осуществляемые умения и навыки.	Фрагментарные знания, частично освоенные умения и навыки.

2. Этапы формирования компетенций и виды оценочных средств

№	Этапы формирования компетенций (разделы дисциплины)	Код и наименование результатов обучения	Вид оценочного средства (тесты, задания, кейсы, вопросы и др.)
1.	Введение в статистический анализ и первичная статистическая обработка	ОР-3.1	Практические задания, теоретические вопросы теста
2.	Критерии сравнения групп	ОР-3.1, ОР-3.2, ОР-3.3	Практические задания, теоретические вопросы теста
3.	Корреляционный анализ	ОР-3.1, ОР-3.2, ОР-3.3	Практические задания, теоретические вопросы теста
4.	Регрессионный анализ	ОР-3.1, ОР-3.2, ОР-3.3	Практические задания, теоретические вопросы теста
5.	Дисперсионный анализ	ОР-3.1, ОР-3.2, ОР-3.3	Практические задания, теоретические вопросы теста
6.	Задачи классификации и кластеризации	ОР-3.1, ОР-3.2, ОР-3.3	Практические задания, теоретические вопросы теста

3. Типовые контрольные задания или иные материалы, необходимые для оценки образовательных результатов обучения

3.1. Типовые задания для проведения текущего контроля успеваемости по дисциплине.

Текущий контроль по дисциплине осуществляется на основании проверки практических заданий, выполняемых студентами на компьютерах в течение семестра.

Студенты получают у преподавателя или собирают самостоятельно статистические данные для дальнейшего анализа и построения математических моделей. Проводят предварительную обработку данных, выбирают адекватный метод анализа в зависимости от целей исследования и типов данных, реализуют выбранные методы на компьютере, делают выводы и интерпретацию полученных результатов.

3.2. Типовые задания для проведения промежуточной аттестации по дисциплине

1. Предварительная обработка данных. Обработка пропущенных значений и выбросов.
2. Критерии проверки нормальности.
3. Параметрические критерии сравнения выборок.
4. Непараметрические критерии сравнения выборок.
5. Общая постановка МНК-оценивания параметров линейной регрессии. Оценивание дисперсии ошибок.
6. Свойства МНК-оценок параметров линейной регрессии.
7. Обобщение оценок параметров линейной регрессии для случая коррелированных гомоскедастичных наблюдений.
8. Обобщение оценок параметров линейной регрессии для случая коррелированных гетероскедастичных наблюдений.
9. Оценки параметров линейной регрессии при связывающих эти параметры ограничениях.
10. Нелинейные модели регрессии, допускающие линеаризацию. Проверка гипотезы об адекватности модели регрессии.
11. Итерационные алгоритмы оценивания параметров регрессии.
12. Доверительные интервалы для параметров регрессии. Интервалы предсказания.
13. Коэффициенты детерминации и парной корреляции, корреляционное отношение: определения и свойства.

14. Частный и множественный коэффициенты корреляции: определения и свойства.
15. Понятие ранговой корреляции. Основные типы задач анализа ранговых связей.
16. Коэффициенты ранговой корреляции Кендалла и Спирмена. Обобщенный коэффициент ранговой корреляции.
17. Проверка гипотезы о статистически зависимой ранговой связи.
18. Коэффициент конкордации и его свойства.
19. Категоризованные данные. Анализ зависимости признаков по таблицам сопряженности.
20. Общая постановка задачи дисперсионного анализа.
21. Однофакторный дисперсионный анализ. Проверка гипотезы о влиянии фактора на исследуемый объект.
22. Исследование влияния на объект уровней фактора методами множественного сравнения.
23. Общее решение задачи двухфакторного дисперсионного анализа.
24. Двухфакторный дисперсионный анализ с равным числом $r > 1$ наблюдений в ячейке.
25. Двухфакторный дисперсионный анализ с неравным числом наблюдений в ячейке.
26. Неполные сбалансированные блоки в задачах дисперсионного анализа.
27. Решение задачи трехфакторного дисперсионного анализа.
28. Общая постановка задачи дискриминантного анализа.
29. Решение задачи параметрического дискриминантного анализа. Расщепление смесей распределений.
30. Типы расстояний и мер близости между объектами и между классами.
31. Типы функционалов качества разбиения множества объектов на классы.
32. Основные типы кластер-процедур.

4. Методические материалы, определяющие процедуры оценивания образовательных результатов обучения

4.1. Методические материалы для оценки текущего контроля успеваемости по дисциплине.

Текущий контроль осуществляется в виде проверки выполнения практических заданий на практических занятиях. Каждая работа оценивается на «зачет»/ «незачет». В течение семестра должны быть в срок выполнены все предложенные работы.

4.2. Методические материалы для проведения промежуточной аттестации по дисциплине.

Экзаменационная оценка складывается из текущего посещения (не менее 80% занятий), в срок выполненных практических заданий и результатов тестирования (при онлайн обучении) или письменного коллоквиума по темам из п. 3.2 (при оффлайн форме).

Посещение и сданные практические задания являются условием для допуска к теоретической части. Оценка за теоретическую часть ставится на основании теста или письменного коллоквиума.

Тест из 15 вопросов. Максимум 30 баллов.

0-15	Неудовлетворительно
16-20	Удовлетворительно
21-25	Хорошо
26-30	Отлично

Письменный коллоквиум. Два вопроса из п. 3.2.

Ответ не дан или дан неверно, имеются грубые ошибки в формулировках и выводах	Неудовлетворительно
Ответ дан, но не в полном объеме, имеются существенные недочеты	Удовлетворительно
Ответ дан практически полностью, имеются некоторые незначительные ошибки	Хорошо
Ответ дан в полном объеме, допускаются очень незначительные погрешности	Отлично

При недостаточном посещении в течение семестра или невыполненных в срок работах студент может получить на экзамене дополнительные вопросы по пропущенным темам или дополнительное задание по практике.