

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ОБРАЗОВАНИЯ

**«Национальный исследовательский ядерный университет «МИФИ»  
Обнинский институт атомной энергетики –**

филиал федерального государственного автономного образовательного учреждения высшего  
профессионального образования «Национальный исследовательский ядерный университет «МИФИ»  
**(ИАТЭ НИЯУ МИФИ)**

**Отделение интеллектуальных кибернетических систем**

Одобрено на заседании УМС  
ИАТЭ НИЯУ МИФИ Протокол от  
30.08.2022 № 1-8/2022

**Методические указания по дисциплине**

**«ДОПОЛНИТЕЛЬНЫЕ ГЛАВЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ И  
МЕТОДОВ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ»**

для студентов направления подготовки

09.04.01 Информатика и вычислительная техника

---

программа:

**Большие данные и машинное обучение в задачах атомной энергетики**

---

Форма обучения: очная

г. Обнинск 2022г.

## 1. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Вид учебного занятия	Организация деятельности студента
Лекция	<p>Написание конспекта лекций: кратко, схематично, последовательно фиксировать основные положения, выводы, формулировки, обобщения; пометить важные мысли, выделять ключевые слова, термины. Проверка терминов, понятий с помощью энциклопедий, словарей, справочников с выписыванием толкований в тетрадь. Обозначить вопросы, термины, материал, который вызывает трудности, пометить и попытаться найти ответ в рекомендуемой литературе. Если самостоятельно не удастся разобраться в материале, необходимо сформулировать вопрос и задать преподавателю на консультации, на практическом занятии и лабораторной работе.</p> <p>Уделить внимание следующим понятиям: Оценивание показателей сложных систем, характеристика статистической информации, понятие о цензурированной выборке, метод максимального правдоподобия, эмпирическое байесовское оценивание, вычисление апостериорной плотности, критерии согласия, параметрические критерии проверки гипотезы об однородности, основные понятия теории планирования эксперимента.</p>
Практические занятия	<p>Работа с конспектом лекций и семинаров, просмотр рекомендуемой литературы. Изучение теоретических основ оценки показателей сложных систем в условиях ограниченной статистической информации, а также критериев проверки статистических гипотез с целью их последующего применения при выполнении курсовой и лабораторных работ.</p>
ИДЗ	<p><b>Этапы выполнения ИДЗ</b></p> <ul style="list-style-type: none"> <li>– Выбор темы и согласование ее с преподавателем в рамках предложенной тематики</li> <li>– Работа с литературой (подбираются источники для написания ИДЗ, изучаются и конспектируются необходимые разделы, относящиеся к теме, записываются библиографические сведения источника и номера страниц, с которых были заимствованы сведения для последующего оформления ссылок на источники.</li> <li>– Составление плана выполнения ИДЗ</li> <li>– Написание основных разделов (подготовленные материалы обрабатывают, анализируют, располагают в соответствии с планом и формируют логические связки между элементами структуры ИДЗ).</li> <li>– Оформление работы (в соответствии с требованиями к оформлению научно-технических отчетов)</li> </ul> <p><b>Структура ИДЗ</b></p> <ul style="list-style-type: none"> <li>– Титульный лист</li> <li>– Оглавление (располагается на следующей после титульного листа странице, представляет собой структуру реферата с указанием наименований разделов и соответствующих им номеров страниц)</li> <li>– Введение (во введении приводят сведения об актуальности темы и степени ее освещенности в литературе, возможно включение и других пунктов).</li> <li>– Основная часть (включает пункты/главы) и подпункты/параграфы, в рамках которых раскрывается тема)</li> <li>– Заключение (содержит краткое изложение основных рассмотренных в реферате вопросов, подведение итогов и выводы)</li> <li>– Список использованной литературы (в список литературы включаются не только цитированные источники, но и литература, изученная при</li> </ul>

	<p>написании работы и упомянутая в тексте).</p> <ul style="list-style-type: none"> <li>– Приложения (при необходимости)</li> </ul> <p>Объем реферата - 10-15 стр.</p> <p><b>Требования к мультимедийной презентации</b></p> <p>Презентация создается с использованием современных компьютерных технологий (например, ПО PowerPoint Microsoft Office)</p> <p>Обязательные структурные элементы презентации: титул (учебное заведение, название темы, автор), введение с обоснованием актуальности темы; слайды, раскрывающие содержание реферата, заключение.</p> <p>Основные требования к содержанию мультимедийной презентации:</p> <ul style="list-style-type: none"> <li>– соответствие содержания презентации поставленным целям и задачам;</li> <li>– соблюдение принятых правил орфографии, пунктуации, сокращений и правил оформления текста;</li> <li>– отсутствие фактических ошибок, достоверность представленной информации;</li> <li>– завершенность (содержание каждой части текстовой информации логически завершено) и связность информации;</li> <li>– лаконичность текста на слайде и при этом максимальная его информативность;</li> <li>– читаемость текста; как правило, кегль шрифта должен быть не менее 24 пунктов;</li> <li>– использование единого стиля оформления; использование для фона слайда спокойного комфортного тона;</li> </ul> <p>Объем презентации – 10-14 слайдов</p>
Контрольная работа	Работа с конспектами лекций, знакомство с основной и дополнительной литературой, включая справочные издания, зарубежные источники, конспект основных положений, терминов, сведений, требующихся для запоминания и являющихся основополагающими в этой теме.
Лабораторная работа	– не предусмотрены
Подготовка к экзамену	При подготовке к экзамену необходимо ориентироваться на конспекты лекций, рекомендуемую литературу и др.

**1. ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ, ВКЛЮЧАЯ ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ (ПРИ НЕОБХОДИМОСТИ)**

- Электронные презентации лекций в формате MS PowerPoint, демонстрируемые с использованием мультимедийного проектора.
- Практические занятия проводятся с объяснением и демонстрацией учебного материала на экране посредством мультимедийного проектора.
- Microsoft Word и PowerPoint для подготовки реферата и презентации.

**2. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ**

- Компьютерный класс сетевых технологий. Класс оснащен 10 компьютерами (Intel Core i5/8GB/1 TB) и 1 компьютером ( Intel Celeron 1.6 GHz, 2 GB RAM, 250 GB) с операционной системой Windows 7, а также мультимедийным проектором. Есть доступ к WI-FI.
- Аудиторный класс, оборудованный проекционным экраном, мультимедийным проектором и персональным компьютером (AMD, ATHLON64, 2.7 GHz, 4 GB RAM, 250 GB). Есть доступ к WI-FI.

## ИНЫЕ СВЕДЕНИЯ И (ИЛИ) МАТЕРИАЛЫ

### 12.1. Перечень образовательных технологий, используемых при осуществлении образовательного процесса по дисциплине

Лекционные занятия проходят с обсуждением учебного материала, демонстрируемого в форме презентаций на экране с использованием мультимедиа-проектора. Значительная часть практических и лабораторных занятий также проводится в интерактивной форме при тесном контакте студентов с преподавателем.

Практические занятия проводятся с использованием современного программного обеспечения, в том числе и авторского. Практические занятия включают в себя элементы деловых игр, исследовательских и аналитических задач. В ходе этих занятий студенты знакомятся с богатым опытом многолетней практической работы по анализу систем на крупных предприятиях.

На первом этапе студенты изучают предметную область в максимально реальной обстановке (по мере возможности). В ходе практических занятий происходит публичное обсуждение результатов обследования. Студенты высказывают свои мнения и дополняют построенную модель.

Каждый студент выполняет три ИДЗ на индивидуальную тему. В этой работе принимают участие все студенты, происходит коллективное обсуждение результатов и итерационное усовершенствование созданного продукта.

### 12.2 Формы организации самостоятельной работы обучающихся (темы, выносимые для самостоятельного изучения; вопросы для самоконтроля; типовые задания для самопроверки

Раздел	Тема	Число часов
1.	<b>Введение в параметрические методы оценки статистической информации</b>	6
1.1.	Оценивание показателей сложных систем и определение их точности.	2
1.2.	Характеристика статистической информации.	4
2.	<b>Метод максимального правдоподобия.</b>	20
2.1.	Оценка параметров законов распределения.	6
2.2.	Обработка цензурированных данных.	8
2.3.	Данные с пропусками, полные и цензурированные наработки	6
3.	<b>Эмпирическое байесовское оценивание.</b>	24
3.1.	Формулировка теоремы Байеса.	4
3.2.	Вычисление апостериорной плотности при последовательном накоплении информации.	6
3.3.	Понятие несобственной плотности распределения . Достаточные статистики. Сопряжённые распределения.	8
3.4.	Формирование априорной плотности распределения.	4

Раздел	Тема	Число часов
3.5.	Примеры байесовского оценивания.	2
<b>4.</b>	<b>Теория проверки статистических гипотез.</b>	<b>16</b>
4.1.	Понятия и определения теории проверки статистических гипотез.	2
4.2.	Критерии согласия. Критерий $\chi^2$ . Критерий А.Н. Колмогорова. Критерий $\omega^2$ .	4
4.3.	Параметрические критерии проверки гипотезы об однородности. Критерий Стьюдента. Критерий Фишера.	4
4.4.	Критерии подобия.	6
<b>5.</b>	<b>Планирование эксперимента.</b>	<b>12</b>
5.1.	Основные понятия теории планирования эксперимента.	4
5.2.	Свойства полного факторного эксперимента.	4
5.3	Дробный факторный эксперимент.	4

Контроль освоения самостоятельно изученного теоретического материала осуществляется в виде собеседования во время защиты лабораторных, в виде устного опроса на практических занятиях и экзамене, на открытом семинаре с презентациями рефератов.

Кроме этого, студенты также самостоятельно выполняют большую часть предусмотренных практических работ, промежуточный результат которых представляется на практических и лабораторных занятиях, а конечный результат - на защите лабораторных и курсовой работ.

#### Вопросы для самоконтроля:

- Оценивание показателей систем, которые являются функционалом от функции распределения:
  - Оценка наработки на отказ,
  - Оценка интенсивности отказов.
  - Оценка вероятности отказа.
- Определение точности показателей систем, являющихся функционалом от функции распределения.
- Дайте определение цензурированной выборки
- Классификация моделей цензурирования.
- Чем характеризуется цензура первого типа.
- Чем характеризуется цензура второго типа.
- Чем характеризуется цензура типа III.
- Какова процедура применения метода максимального правдоподобия для оценивания параметров законов распределения.
- Понятие функции правдоподобия.
- Логарифмическая функция правдоподобия.
- Сформируйте функцию правдоподобия для выборок, содержащих полные данные.
- Сформируйте функцию правдоподобия для выборок, содержащих цензурированные наработки.
- Сформируйте функцию правдоподобия для выборок, содержащих данные с пропусками.
- Как формируется информационная матрица Фишера.
- Как определяется точность оценок параметров в методе максимального правдоподобия

16. Оценка вероятностных показателей систем путём обработки цензурированных справа данных.
17. Оценка вероятностных показателей систем путём обработки цензурированных слева данных.
18. Оценивание показателей систем по группированным данным.
19. Сформулируйте теорему Байеса для событий.
20. Сформулируйте теорему Байеса для непрерывных случайных величин.
21. Как формируется апостериорная плотность распределения при последовательном накоплении информации.
22. Понятие несобственной плотности распределения.
23. Что такое достаточные статистики.
24. Дайте определение сопряжённых распределений.
25. Какие методы применяются для формирования априорной плотности распределения оцениваемого параметра.
26. Пример оценивания параметров нормального закона распределения.
27. Пример оценивания параметров семейства гамма-распределений.
28. Байесовское оценивание параметров по многократно цензурированным данным.
29. Как формируется функции риска.
30. Порядок вычисления оценок характеристик законов распределения методом минимизации функции риска.
31. Как определяется точность оцениваемых показателей в методе минимизации функции риска.
32. Схема байесовского оценивания в случае, когда наблюдения за функционированием объектов проводятся в несколько этапов.
33. Способы формирования апостериорного распределения.
34. Несобственная плотность распределения в случае биномиального распределения.
35. Несобственная плотность распределения в случае нормального закона распределения.
36. Понятие характеристической функции.
37. Преобразование Фурье, взаимосвязь плотности распределения и характеристической функции.
38. Решающее правило. Основная и конкурирующие гипотезы.
39. Проверка гипотез. Ошибки первого и второго рода.
40. Проверка гипотез. Мощность критерия.
41. Проверка гипотез. Построение доверительных границ.
42. Способы построения статистических критериев проверки гипотезы об однородности.
43. Изложите суть критерия Стьюдента.
44. Изложите суть критерия Фишера.
45. Изложите суть критерия подобия.
46. Критерий подобия. Вычисление критических значений принятия гипотезы для нормального закона.
47. Критерий подобия. Вычисление критических значений принятия гипотезы для гамма-распределения.
48. Вычисление мощности критерия.

### 12.3. Краткий терминологический словарь

Априорная плотность распределения	Распределение оцениваемого показателя до проведения наблюдения за случайной величиной, по реализации которой производится оценка показателя
Апостериорная плотность распределения	Распределение оцениваемого показателя после проведения наблюдения за случайной величиной, по реализации которой производится оценка показателя
Границы критической	Точки, отделяющие критическую область от области принятия

области	гипотезы
Критическая область	Совокупность значений статистики критерия, при которых нулевую гипотезу отвергают
Мощность критерия	Вероятность попадания критерия в критическую область при условии, что справедлива конкурирующая гипотеза
Область принятия гипотезы	Совокупность значений статистики критерия, при которых нулевую гипотезу принимают
Ошибка первого рода	Вероятность отвергнуть гипотезу при условии, что она верна
Ошибка второго рода	Вероятность принятия неправильной гипотезы
Оценка максимального правдоподобия	Точечная оценка параметра, при которой функция правдоподобия достигает максимума
Плотность распределения	Функция, которая является первой производной от функции распределения
Случайная величина $X$	Вещественная $A$ -измеримая функция, значение $X(s)$ которой определено для каждой точки $s$ принадлежащей $S$
Случайное событие	Событие, которое при осуществлении совокупности условий может реализоваться, а может не реализоваться
Статистический критерий	Случайная величина, которая служит для проверки нулевой гипотезы
Точность оценки	Характеризуется границами доверительного интервала построенного для оцениваемого показателя
Функция правдоподобия	Произведение значений плотностей распределения наблюдаемой случайной величины, которые вычисляются в точках, зафиксированных в результате наблюдений
Функция риска	Характеристика среднего ущерба, которое несет лицо принимающее решение, в результате конкретного выбора