

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Казанский национальный исследовательский технологический
университет»
(ФГБОУ ВО «КНИТУ»)

УТВЕРЖДАЮ
Проректор по УР
Бурмистров А.В.

« 1 » 2019 г.

РАБОЧАЯ ПРОГРАММА

По дисциплине «Вычислительная статистика»

Направление подготовки - 02.03.03 «Математическое обеспечение и администрирование информационных систем»

Профиль/специализация – Информационные системы и базы данных

Квалификация выпускника - бакалавр

Форма обучения - очная

Институт, факультет - Нефти, химии и нанотехнологий, Наноматериалов и нанотехнологий

Кафедра-разработчик рабочей программы - Интеллектуальных систем и управления информационными ресурсами

Курс 3, семестр 5

	Часы	Зачетные единицы
Лекции	18	0,5
Практические занятия		
Лабораторные занятия	36	1
Контроль самостоятельной работы		
Самостоятельная работа	63	1,25
Форма аттестации экзамен	27	0,75
Всего	144	4

Казань, 2019 г.

Л.А. Китаева

1. Цели освоения дисциплины

Целями освоения дисциплины «Вычислительная статистика» являются

- а) формирование знаний о вычислительных и статических методах обработки данных,
- б) обучение технологии расчета статистических характеристик,
- в) обучение способам применения программного обеспечения для статической обработки данных,
- г) раскрытие сущности случайных процессов, происходящих в системах

2. Место дисциплины (модуля) в структуре основной образовательной программы

Дисциплина «Вычислительная статистика» относится к части ООП формируемой участниками образовательных отношений и формирует у бакалавров по направлению подготовки 02.03.03 набор знаний, умений, навыков и компетенций.

Для успешного освоения дисциплины «Вычислительная статистика» бакалавр по направлению подготовки 02.03.03 должен освоить материал предшествующих дисциплин должен освоить материал предшествующих дисциплин:

- а) математический анализ
- б) линейная алгебра
- в) теория вероятностей и математическая статистика

Дисциплина «Вычислительная статистика» является предшествующей и необходима для успешного усвоения последующих дисциплин:

- а) современные методы принятия решений
- б) распределенная обработка и анализ больших данных
- в) информационный поиск и обработка текстов на естественном языке

Знания, полученные при изучении дисциплины могут быть использованы при прохождении практик и выполнении выпускной квалификационной работы.

3. Компетенции и индикаторы достижения компетенции обучающегося, формируемые в результате освоения дисциплины

ПК-5 Способен использовать современные методы разработки и реализации конкретных алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования

ПК-5.1 - Знает современные методы разработки и реализации алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования

ПК-5.2 – Умеет разрабатывать и реализовывать алгоритмы математических моделей на базе языков программирования и пакетов прикладных программ моделирования

ПК-5.3 - Владеет навыками разработки и реализации алгоритмов на базе языков программирования и пакетов прикладных программ моделирования

В результате освоения дисциплины обучающийся должен

1) Знать:

- а) основные методы математической статистики,
- б) законы распределения случайных величин.

2) Уметь:

- а) применять методы математической статистики для решения вероятностных и статистических задач,
- б) пользоваться математической литературой для самостоятельного изучения теоретических и практических задач и вопросов.

3) Владеть:

- а) методами решения алгебраических уравнений,
- б) задач дифференциального и интегрального исчисления,
- в) методами построения вероятностных и статистических моделей для задач, связанных с профессиональной деятельностью.

4. Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 4 зачетные единицы, 144 часов.

№ п /п	Раздел дисциплины	Семестр	Виды учебной работы (в часах)					Оценочные средства для проведения промежуточной аттестации по разделам
			Лекции	Практиче ские занятия	Лаборат орные работы	КСР	СРС	
1	Моделирование случайных величин и векторов. Методы Монте-Карло.	5	3		6		11	
2	Марковские цепи. МСМС. Алгоритм имитации отжига.	5	3		6		11	Контрольная работа
3	ЕМ алгоритм	5	3		6		11	Контрольная работа
4	Ресемплинг. Метод складного ножа и бутстреп.	5	3		6		10	Контрольная работа
5	Ресемплинг. Перестановочные тесты для проверки различных статистических гипотез.	5	3		6		10	
6	Проверка гипотез с помощью бутстрепа. Проверка гипотез с помощью симуляций	5	3		6		10	Контрольная работа
ИТОГО			18		36		63	
Форма аттестации					Экзамен (27ч.)			

5. Содержание лекционных занятий по темам с указанием формируемых компетенций

№	Раздел дисциплины	Часы	Тема лекционного занятия, краткое содержание	Индикаторы достижения компетенции
1	Моделирование случайных величин и векторов. Методы Монте-Карло	3	Генераторы случайных чисел. Генерирование равномерного распределения. Метод обратного преобразования. Генерирование дискретных случайных величин. Генерирование абсолютно непрерывных случайных величин. Метод выборки с отклонением. Генерирование стандартного нормального распределения и двумерного Гауссовского вектора. Проблема “проклятья размерности” в этих задачах. Генерирование смеси распределений. Моделирование случайных векторов через условное распределение. Оценивание параметров и интегрирование методом Монте-Карло. Выборка по значимости (importance sampling).	ПК-5.1, ПК-5.2, ПК-5.3
2	Марковские цепи. МСМС. Алгоритм имитации отжига.	3	Переходные матрицы и плотности вероятностей Марковских цепей. Стационарное распределение Марковской цепи. Уравнение баланса. Монте-Карло с помощью Марковских цепей (МСМС). Алгоритм Метрополис-Гастингса. Моделирование многомерных распределений. Семплинг по Гиббсу. Модель Поттса (в частности, модель Изинга) и ее моделирование.	ПК-5.1, ПК-5.2, ПК-5.3

			Случайные графы Эрдоша-Реньи и их моделирование. Алгоритм имитации отжига для ускорения МСМС и в применении к оптимизационным задачам. Задача коммивояжера. Стохастический градиентный спуск в сравнении с МСМС.	
3	ЕМ алгоритм	3	Модель смеси Гауссовских распределений. Вероятностный метод кластеризации в такой модели. Сравнение ЕМ и K-means алгоритмов. Применение ЕМ алгоритма к смеси Марковских цепей.	ПК-5.1, ПК-5.2, ПК-5.3
4	Ресемплинг. Метод складного ножа и бутстреп.	3	Улучшение и объединение оценок. Идеальный бутстреп. Бутстреп для оценки среднего, дисперсии и других параметров выборки. Доверительные интервалы параметров с помощью бутстрепа. Свойства получаемых оценок. Коррекция смещения бутстрепом. Сравнение метода складного ножа и бутстрепа. Их асимптотические свойства. Bootstrap aggregating (bagging). Boosting.	ПК-5.1, ПК-5.2, ПК-5.3
5	Ресемплинг. Перестановочные тесты для проверки различных статистических гипотез.	3	Нахождение р-значения перестановками из исходной выборки. Перестановочные тесты на однородность. Задача валидация модели. Перестановочные тесты для проверки независимости. Точный тест Фишера.	ПК-5.1, ПК-5.2, ПК-5.3
6	Проверка гипотез с помощью бутстрепа. Проверка гипотез с помощью симуляций	3	Нахождение р-значения симуляциями, бутстрепом. Точный тест для пропорций.	ПК-5.1, ПК-5.2, ПК-5.3

6. Содержание практических занятий

Не предусмотрено учебным планом.

7. Содержание лабораторных занятий

Целью проведения лабораторных работ является закрепление теоретического материала на наглядном примере, а также приобретение практических навыков системного администрирования.

№ п/п	Раздел дисциплины	Часы	Наименование лабораторной работы	Индикаторы достижения компетенции
1	Моделирование случайных величин и векторов. Методы Монте-Карло.	6	Генерирование дискретных случайных величин. Генерирование стандартного нормального распределения и двумерного Гауссовского вектора.	ПК-5.1, ПК-5.2, ПК-5.3
2	Марковские цепи. МСМС. Алгоритм имитации отжига.	6	Моделирование многомерных распределений. Семплинг по Гиббсу.	ПК-5.1, ПК-5.2, ПК-5.3
3	ЕМ алгоритм	6	Применение ЕМ алгоритма к смеси Марковских цепей.	ПК-5.1, ПК-5.2, ПК-5.3
4	Ресемплинг. Метод складного ножа и бутстреп.	6	Сравнение метода складного ножа и бутстрепа. Их асимптотические свойства. Bootstrap aggregating (bagging). Boosting.	ПК-5.1, ПК-5.2, ПК-5.3
5	Ресемплинг. Перестановочные тесты для проверки различных статистических гипотез.	6	Нахождение р-значения перестановками из исходной выборки. Перестановочные тесты на однородность.	ПК-5.1, ПК-5.2, ПК-5.3
6	Проверка гипотез с помощью бутстрепа. Проверка гипотез с помощью симуляций	6	Нахождение р-значения симуляциями, бутстрепом. Точный тест для пропорций.	ПК-5.1, ПК-5.2, ПК-5.3

8. Самостоятельная работа

№ п/п	Темы, выносимые на самостоятельную работу	Часы	Форма СРС	Индикаторы достижения компетенции
1	Моделирование случайных величин и векторов. Методы Монте-Карло.	11	Проработка теоретического материала, подготовка к лабораторной работе	ПК-5.1, ПК-5.2, ПК-5.3
2	Марковские цепи. МСМС. Алгоритм имитации отжига.	11	Проработка теоретического материала, подготовка к лабораторной работе	ПК-5.1, ПК-5.2, ПК-5.3
3	ЕМ алгоритм	11	Проработка	ПК-5.1,

			теоретического материала, подготовка к лабораторной работе	ПК-5.2, ПК-5.3
4	Ресемплинг. Метод складного ножа и бутстреп.	10	Проработка теоретического материала, подготовка к лабораторной работе	ПК-5.1, ПК-5.2, ПК-5.3
5	Ресемплинг. Перестановочные тесты для проверки различных статистических гипотез.	10	Проработка теоретического материала, подготовка к лабораторной работе	ПК-5.1, ПК-5.2, ПК-5.3
6	Проверка гипотез с помощью бутстрепа. Проверка гипотез с помощью симуляций	10	Проработка теоретического материала, подготовка к лабораторной работе	ПК-5.1, ПК-5.2, ПК-5.3

9. Использование рейтинговой системы оценки знаний

При оценке результатов деятельности, обучающихся в рамках дисциплины «Вычислительная статистика» используется рейтинговая система. Рейтинговая оценка формируется на основании текущего и промежуточного контроля. Максимальное и минимальное количество баллов по различным видам учебной работы описано в «Положении о балльно-рейтинговой системе оценки знаний студентов и обеспечения качества учебного процесса» ФГБОУ ВО КНИТУ.

При изучении указанной дисциплины предусматривается сдача четырех контрольных работ с максимальным количеством баллов 10 за каждый и сдача шести лабораторных работ с максимальным количеством баллов 10 за каждый.

В итоге максимальный рейтинг за изучение дисциплины составляет 100 баллов за семестр.

<i>Оценочные средства</i>	<i>Кол-во</i>	<i>Min, баллов</i>	<i>Max, баллов</i>
<i>Лабораторная работа</i>	<i>6</i>	<i>36</i>	<i>60</i>
<i>Контрольная работа</i>	<i>4</i>	<i>24</i>	<i>40</i>
<i>Итого:</i>		<i>60</i>	<i>100</i>

10. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины

Оценочные средства для проведения текущего контроля успеваемости, промежуточной аттестации обучающихся и итоговой (государственной итоговой) аттестации разрабатываются согласно положению о Фондах оценочных средств, рассматриваются как составная часть рабочей программы и оформляются отдельным документом.

11. Информационно-методическое обеспечение дисциплины

11.1. Основная литература

При изучении дисциплины «Вычислительная статистика» в качестве основных источников информации рекомендуется использовать следующую литературу.

Основные источники информации	Кол-во экз.
Математика и информатика: практикум / Е.Н. Гусева, И.Ю. Ефимова, Р.И. Коробков и др. – 4-е изд., стер. – Москва: Флинта, 2016. – 399 с.	ЭБС «Университетская библиотека онлайн»: http://biblioclub.ru/index.php?page=book_red&id=83437 доступ после регистрации с IP-адресов КНИТУ
Нагаев, В.В. Информатика и математика: учебное пособие / В.В. Нагаев, В.Н. Сотников, А.М. Попов; ред. А.М. Попов. – Москва: Юнити, 2015. – 302 с. ISBN 978-5-238-01396-1	ЭБС «Университетская библиотека онлайн»: http://biblioclub.ru/index.php?page=book_red&id=436808 доступ после регистрации с IP-адресов КНИТУ

11.2. Дополнительная литература

В качестве дополнительных источников информации рекомендуется использовать следующую литературу:

Дополнительные источники информации	Кол-во экз.
Интеллектуальные и информационные системы в медицине: мониторинг и поддержка принятия решений. – Москва; Берлин: Директ-Медиа, 2016. – 529 с. ISBN 978-5-4475-7150-4	ЭБС «Университетская библиотека онлайн»: http://biblioclub.ru/index.php?page=book_red&id=434736 доступ после регистрации с IP-адресов КНИТУ

11.3. Электронные источники информации

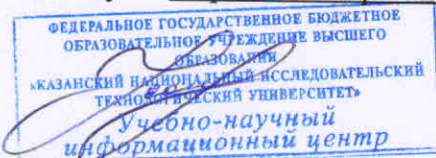
При изучении дисциплины «Вычислительная статистика» в качестве электронных источников информации, рекомендуется использовать следующие источники:

Электронный каталог УНИЦ КНИТУ – режим доступа: <http://ruslan.kstu.ru/>

ЭБС «Университетская библиотека онлайн» -режим доступа <http://biblioclub.ru>

ЭБС «IPRBooks» -режим доступа <http://www.iprbookshop.ru>

Согласовано:
Зав.сектором ОКУФ



11.4. Современные профессиональные базы данных и информационные справочные системы.

1. eLIBRARY.ru - крупнейший российский информационный портал в области науки, технологии, медицины и образования. Доступ свободный: www.elibrary.ru

документов, из более 3 000 журналов и 170 000 книг по математике, статистике, информатике, а также машиностроению, физике, естественным наукам и др. Доступ свободный: zbmath.org

3. Архив журналов РАН. Доступ свободный: elibrary.ru и libnauka.ru

12. Материально-техническое обеспечение дисциплины (модуля).

Учебные аудитории для проведения учебных занятий оснащены оборудованием:

1. smart-доска

Помещения для самостоятельной работы оснащены компьютерной техникой:

1. персональный компьютер с возможностью подключения к сети «Интернет» и обеспечены доступом в электронную информационную среду КНИТУ. Допускается замена оборудования его виртуальными аналогами.

13. Образовательные технологии

Из общего количества часов 4 проводится в интерактивной форме. Интерактивные занятия реализуются с помощью дискуссий и лекций-дискуссий.