

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 № 2-8/2022

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

ОБРАБОТКА И СТАТИСТИЧЕСКИЙ АНАЛИЗ БОЛЬШИХ ДАННЫХ

для студентов направления подготовки

09.04.01 – Информатика и вычислительная техника

программа

Большие данные и машинное обучение в атомной энергетике

Форма обучения: очная

г. Обнинск 2022г.

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

Целью освоения учебной дисциплины «Статистический анализ данных» является обучение студентов основным непараметрическим статистическим методам обработки экспериментальных данных, закрепление знаний, полученных в курсе «Теория вероятностей, математическая статистика и случайные процессы», углубление знаний, полученных в курсах «Теория вероятностей, математическая статистика и случайные процессы», «Статистические методы и модели в управлении», развитие исследовательских навыков работы.

В результате освоения ООП магистратуры обучающийся должен овладеть следующими результатами обучения по дисциплине:

Коды компетенций	Результаты освоения ООП <i>Содержание компетенций</i>	Перечень планируемых результатов обучения по дисциплине
ПК-1	Способен применять научно обоснованные перспективные методы исследования и решать задачи на основе знания мировых тенденций развития вычислительной техники и информационных технологий с внедрением результатов исследований в реальный сектор экономики	Знать: основные типы (модели) распределений вероятности, основные статистические методы параметрического и непараметрического оценивания распределения случайной величины, основные параметрические и непараметрические методы проверки статистических гипотез. Уметь: правильно выбирать структуру и формат данных; корректно заносить статистическую информацию; обрабатывать эту информацию; разрабатывать математические модели исследуемых процессов и изделий; выбирать методики и средства решения задачи. Владеть: методами обработки экспериментальных данных с помощью ЭВМ; подготовкой научно-технических отчетов, обзоров, публикаций по результатам выполненных исследований.
СПК-1	Способен использовать и развивать методы научных исследований и инструментарий в области интеллектуального анализа данных	Знать: различные способы хранения статистической информации на компьютере; возможности компьютерной обработки данных; методы сбора, обработки, анализа и систематизации научно-технической информации по теме исследования. Уметь: ставить и решать задачи статистической обработки, анализировать имеющуюся экспериментальную информацию с различных точек зрения и строить адекватную вероятностную модель, корректно выбирать необходимую статистическую процедуру. Владеть: методами обработки экспериментальных данных, анализа распределений и зависимостей.

2. Место дисциплины в структуре ООП магистратуры

Дисциплина реализуется в рамках базовой части.

Для освоения дисциплины необходимы компетенции, сформированные в рамках изучения следующих дисциплин бакалавриата: «Математический анализ», «Алгебра и геометрия», «Дискретная математика», «Математическая логика и теория алгоритмов», «Теория вероятностей, математическая статистика и случайные процессы», «Статистические методы и модели в управлении».

Дисциплина изучается на 1 курсе во 2 семестре.

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

Общая трудоемкость (объем) дисциплины составляет 4 зачетных единицы (з.е.), 144 академических часа.

3.1. Объём дисциплины по видам учебных занятий (в часах)

Объем дисциплины	Всего часов
	Очная форма обучения
Общая трудоемкость дисциплины	144
Контактная работа обучающихся с преподавателем (по видам учебных занятий) (всего)	32
Аудиторная работа (всего):	32
<i>Лекции</i>	16
<i>семинары, практические занятия</i>	16
<i>лабораторные работы</i>	16
Внеаудиторная работа (всего):	–
<i>индивидуальная работа обучающихся с преподавателем:</i>	–
курсовое проектирование	–
групповая, индивидуальная консультация и иные виды учебной деятельности, предусматривающие групповую или индивидуальную работу обучающихся с преподавателем)	–
Творческая работа (эссе)	–
Самостоятельная работа обучающихся (всего)	60
Вид промежуточной аттестации обучающегося: экзамен	36

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

4.1. Разделы дисциплины и трудоемкость по видам учебных занятий (в академических часах)

№ п/п	Наименование раздела /темы дисциплины	Общая трудоёмкость всего (в часах)	Виды учебных занятий, включая самостоятельную работу обучающихся и трудоемкость (в часах)				СРО	Формы текущего контроля успеваемости
			Аудиторные учебные занятия*			СРО		
			Лек	Сем/Пр	Лаб			
1	Анализ распределения	11	6	5	–	25		
1.1	Оценка функции распределения	4	2	2	–	9	Контрольная работа № 1	
1.2	Оценка плотности	4	2	2	–	8	Контрольная	

	распределения						работа № 1
1.3	Порядковые статистики	3	2	1	–	8	Реферат
2.	Анализ зависимости	11	6	5	–	25	
2.1	Ранги и ранговая корреляция	5	3	2	–	15	Контрольная работа № 2
2.2	Непараметрический регрессионный анализ	6	3	3	–	10	Реферат
3.	Непараметрические критерии	10	5	5	–	26	
3.1	Задача о положении	2	1	1	–	8	Контрольная работа № 1, реферат
3.2	Задача о рассеянии	2	1	1	–	8	Контрольная работа № 1, реферат
3.3	Дисперсионный анализ	4	2	2		6	Контрольная работа № 2, реферат
3.4	Другие критерии и задачи	2	1	1		4	Реферат

4.2. Содержание дисциплины, структурированное по разделам (темам)

4.2.1. Лекционный курс

№	Наименование раздела /темы дисциплины	Содержание
1	Анализ распределения	
1.1	Оценка функции распределения	Эмпирическая функция распределения и ее свойства. Методы сглаживания эмпирической функции распределения. Анализ влияния параметра сглаживания на оценку. Гипотеза согласия. Теоремы Колмогорова. Теорема Гливленко-Кантелли
1.2	Оценка плотности распределения	Гистограмма и полигон частот- как простейшие непараметрические методы оценивания плотности и их свойства. Оптимальное разбиение. Ядерная оценка плотности. Ядро и параметр сглаживания. Проекционная оценка плотности. Оптимальное число гармоник.
1.3	Порядковые статистики	Ранг и порядковые статистики. Распределение порядковых статистик. Непараметрические доверительные интервалы для квантилей. Непараметрические толерантные интервалы
2.	Анализ зависимости	
2.1	Ранги и ранговая корреляция	Ковариация и корреляция и их свойства (повтор). Выборочная ковариация и корреляция. Коэффициенты ранговой корреляции Спирмена и Кендалла. Корреляционный и регрессионный анализ.
2.2	Непараметрический регрессионный анализ	Линейная регрессия и метод наименьших квадратов. Метод локальной аппроксимации. Параметр локальности. Сходимость оценок регрессии. Чувствительность оценок к неоднородности наблюдений. Множественная линейная регрессия, частные коэффициенты корреляции, их смысл и оценка по выборке. Задача о регрессии и угле наклона.
3.	Непараметрические критерии	

3.1	Задача о положении	Задача о дихотомических данных. Одновыборочная задача о положении. Двухвыборочная задача о положении. Критерии Фишера, Вилкоксона и т.д. и оценки, связанные с ними.
3.2	Задача о рассеянии	Двухвыборочная задача о рассеянии. Критерии Ансари-Бредли, Мозеса и т.д. и оценки, связанные с ними.
3.3	Дисперсионный анализ	Однофакторные таблицы дисперсионного анализа. Критерии Краскела-Уоллиса, Джонкхиера-Терпстры и оценки, связанные с ними. Двухфакторные таблицы дисперсионного анализа. Критерии Фридмана, Пейджа и оценки, связанные с ними.
3.4	Другие критерии и задачи	Критерии для произвольных альтернатив. Критерии экспоненциальности, нормальности. Критерии случайности и т.д.

4.2.2. Практические/семинарские и лабораторные занятия

№	Наименование раздела /темы дисциплины	Содержание
1	Анализ распределения	
1.1	Оценка функции распределения	Эмпирическая функция распределения и ее свойства. Нахождение статистик Колмогорова, Мизеса. Анализ влияния параметра сглаживания на оценку.
1.2	Оценка плотности распределения	Гистограмма и полигон частот. Оптимальное разбиение. Построение ядерной оценки плотности. Ядро и параметр сглаживания. Построение проекционной оценки плотности. Оптимальное число гармоник.
1.3	Порядковые статистики	Ранг и порядковые статистики. Распределение порядковых статистик. Непараметрические доверительные интервалы для квантилей. Непараметрические толерантные интервалы
2.	Анализ зависимости	
2.1	Ранги и ранговая корреляция	Средний ранг. Дисперсия ранга. Коэффициенты ранговой корреляции Спирмена и Кендалла. Математическое ожидание и дисперсия, распределение коэффициентов ранговой корреляции.
2.2	Непараметрический регрессионный анализ	Линейная регрессия и метод наименьших квадратов. Метод локальной аппроксимации. Параметр локальности. Множественная линейная регрессия, частные коэффициенты корреляции, их смысл и оценка по выборке. Задача о регрессии и угле наклона. Оценка Тейла.
3.	Непараметрические критерии	
3.1	Задача о положении	Задача о дихотомических данных. Одновыборочная задача о положении. Двухвыборочная задача о положении. Статистики Фишера, Вилкоксона и т.д. и их характеристики.
3.2	Задача о рассеянии	Двухвыборочная задача о рассеянии. Статистики Ансари-Бредли, Мозеса и т.д. и их характеристики.
3.3	Дисперсионный анализ	Однофакторные таблицы дисперсионного анализа. Статистики Краскела-Уоллиса, Джонкхиера-Терпстры и их характеристики. Двухфакторные таблицы дисперсионного анализа. Статистики Фридмана, Пейджа и их характеристики.
3.4	Другие критерии и задачи	Критерии для произвольных альтернатив. Критерии экспоненциальности, нормальности. Критерии случайности

		и т.д.
--	--	--------

4.2.3. Курсовая работа – не предусмотрена

5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

В качестве учебно-методических материалов используется рекомендованная литература

6. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

6.1. Паспорт фонда оценочных средств по дисциплине

№ п/п	Контролируемые разделы (темы) дисциплины (результаты по разделам)	Код контролируемой компетенции (или её части) / и ее формулировка	Наименование оценочного средства
1.1	Оценка функции распределения	ПК-1: способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Контрольная работа № 1 (решение задач)
		СПК-1: умением осуществлять моделирование процессов и объектов на базе стандартных пакетов автоматизированного проектирования и исследований	Контрольная работа № 1 (решение задач)
1.2	Оценка плотности распределения	ПК-1: способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Контрольная работа № 1 (решение задач)
		ПК-1: умением осуществлять моделирование процессов и объектов на базе стандартных пакетов автоматизированного проектирования и исследований	Контрольная работа № 1 (решение задач)
1.3	Порядковые статистики	СПК-1: способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания,	Реферат

		умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	
		СПК-10: умением осуществлять моделирование процессов и объектов на базе стандартных пакетов автоматизированного проектирования и исследований	Реферат
2.1	Ранги и ранговая корреляция	ПК-1: способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Контрольная работа № 2 (решение задач)
		СПК-1: умением осуществлять моделирование процессов и объектов на базе стандартных пакетов автоматизированного проектирования и исследований	Контрольная работа № 2 (решение задач)
2.2	Непараметрический регрессионный анализ	ПК-1: способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Реферат
		СПК-1: умением осуществлять моделирование процессов и объектов на базе стандартных пакетов автоматизированного проектирования и исследований	Реферат
3.1	Задача о положении	ПК-1: способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения	Контрольная работа № 1 (решение задач)

		нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	
		СПК-1: умением осуществлять моделирование процессов и объектов на базе стандартных пакетов автоматизированного проектирования и исследований	Контрольная работа № 1 (решение задач)
		ПК-1: способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Реферат
		СПК-1: умением осуществлять моделирование процессов и объектов на базе стандартных пакетов автоматизированного проектирования и исследований	Реферат
3.2	Задача о рассеянии	СПК-1: способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Контрольная работа № 1 (решение задач)
		СПК-1: умением осуществлять моделирование процессов и объектов на базе стандартных пакетов автоматизированного проектирования и исследований	Контрольная работа № 1 (решение задач)
		ПК-1: способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном	Реферат

		контексте	
		СПК-1: умением осуществлять моделирование процессов и объектов на базе стандартных пакетов автоматизированного проектирования и исследований	Реферат
3.3	Дисперсионный анализ	ПК-1: способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Контрольная работа № 2 (решение задач)
		СПК-1: умением осуществлять моделирование процессов и объектов на базе стандартных пакетов автоматизированного проектирования и исследований	Контрольная работа № 2 (решение задач)
		ПК-1: способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Реферат
		СПК-1: умением осуществлять моделирование процессов и объектов на базе стандартных пакетов автоматизированного проектирования и исследований	Реферат
3.4	Другие критерии и задачи	ПК-1: способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Реферат

		СПК-1: умением осуществлять моделирование процессов и объектов на базе стандартных пакетов автоматизированного проектирования и исследований	Реферат
--	--	--	---------

6.2. Типовые контрольные задания или иные материалы

6.2.1. Экзамен

Экзамен проводится в письменной форме и состоит в решении ряда практических задач непараметрической статистики. При решении этих задач используются следующие разделы теории:

1. Эмпирическая функция распределения и ее свойства.
2. Гистограмма и ее свойства.
3. Ядерная и проекционная оценка и ее свойства.
4. Индикаторная функция.
5. Ранг, ранговая статистика.
6. Ранговые корреляции.
7. Построение рангового критерия.

Три типовых билета

Письменный экзамен по курсу
«Дополнительные главы прикладной математики».
Билет №1

1. Средний балл весенних сессий некоторых групп ИАТЭ:

	BT1	BT2	A1	ИС
I курс	3.5	3.3	4.5	3.3
II курс	3.3	3.3	4.9	3.4
III курс	4.4	3.2	5	4.1
IV курс	4.2	4.4	5	4.2
V курс	4.8	5	5	4.3

Можно ли утверждать, что по мере взросления студенты лучше учатся? Проверить нулевую гипотезу об отсутствии такого тренда с уровнем значимости $\alpha_0 = 0.05$ в предположении отсутствия фактора– группа.

2. Найти с помощью эмпирической ФР $F_n^{\downarrow}(x)$, построенной по выборке $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ эмпирическую ФР случайной величины $|\xi| - 2$.
3. Найти с помощью эмпирической ФР $F_n^{\downarrow}(x)$, построенной по выборке $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ вероятность события: $P(\xi^2 - \xi < 0)$.
4. Найти вероятность $P(nF_n^{\downarrow}(x) = 0.5)$.

5. Пусть $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ - выборка. $f_{\xi}(x)$ - оцениваемая плотность распределения вероятностей. $\Delta_1, \Delta_2, \dots, \Delta_s$ - конечное разбиение. Найти $Mv\{\Delta_1\}$, где $v\{A\}$ - частота заполнения множества A .
6. $f_n^{\square}(x)$ - гистограммная оценка плотности. Найти $\int_{-\infty}^{\infty} f_n^{\square}(2x-3) dx$.
7. Чему равен коэффициент ранговой корреляции Спирмена для двух наборов данных: $\vec{x} = (x_1, x_2, \dots, x_n)$ и $\vec{y} = (y_1, y_2, \dots, y_n)$, если $x_i = 2y_i$ для всех $i = 1, 2, \dots, n$?
8. Для данных: $\vec{x} = (1, 2, 3, 5, 7, 1)$, $\vec{y} = (0, 1, 3, 1)$ посчитать ранг $r_3^{(x)}$. Если для проверки гипотезы об отсутствии сдвига $H_0: \Delta = Mx - My = 0$ против альтернативы $H_1^+: \Delta > 0$ применяется статистика $W = \sum_{i=1}^n (r_i^{(x)})^2$, то где должна находиться критическая область слева или справа? Ответ обоснуйте.

Письменный экзамен по курсу
«Дополнительные главы прикладной математики».
Билет №2

1. Средний балл весенних сессий некоторых групп ИАТЭ:

	BT1	BT2	A1	ИС
I курс	3.5	3.3	4.5	3.3
II курс	3.3	3.3	4.9	3.4
III курс	4.4	3.2	5	4.1
IV курс	4.2	4.4	5	4.2
V курс	4.8	5	5	4.3

Можно ли утверждать, что по мере взросления студенты лучше учатся? Проверить нулевую гипотезу об отсутствии такого тренда с уровнем значимости $\alpha_0 = 0.05$ в предположении наличия фактора – группа.

2. Найти с помощью эмпирической ФР $F_n^{\square}(x)$, построенной по выборке $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ эмпирическую ФР случайной величины $2\xi - 5$.
3. Найти с помощью эмпирической ФР $F_n^{\square}(x)$, построенной по выборке $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ вероятность события: $P(\ln|\xi| < 1)$.
4. Найти вероятность $P(10F_{10}^{\square}(x) = 2)$.
5. Пусть $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ - выборка без связей, $\vec{r} = (r_1, r_2, \dots, r_n)$ - ее ранги. Найти $M(r_1 \cdot r_2)$.
6. Пусть $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ - выборка. $f_{\xi}(x)$ - оцениваемая плотность распределения вероятностей. $\Delta_1, \Delta_2, \dots, \Delta_s$ - конечное разбиение. Найти $Mv\{\Delta_1 \cap \Delta_2\}$, где $v\{A\}$ - частота заполнения множества A .
7. Чему равен коэффициент ранговой корреляции Спирмена для двух наборов данных: $\vec{x} = (x_1, x_2, \dots, x_n)$ и $\vec{y} = (y_1, y_2, \dots, y_n)$, если $x_i = 3 - 2y_i$ для всех $i = 1, 2, \dots, n$?

8. Для данных: $\vec{x} = (1, 2, 3, 5, 7, 1)$, $\vec{y} = (0, 1, 3, 1)$ посчитать ранг $r_3^{(y)}$. Если для проверки гипотезы об отсутствии сдвига $H_0: \Delta = Mx - My = 0$ против альтернативы $H_1^+: \Delta > 0$

$$W = \sum_{i=1}^n r_i^{(y)}$$

применяется статистика W , то где должна находиться критическая область слева или справа? Ответ обоснуйте.

Письменный экзамен по курсу
«Дополнительные главы прикладной математики».
Билет №3

1. Концентрация студентов в аудитории по месяцам в различных аудиториях ИАТЭ в процентах:

	сентябрь	октябрь	ноябрь	декабрь
I курс	99	98	28	100
II курс	88	50	40	95
III курс	44	44	33	90
IV курс	42	52	62	72
V курс	52	63	73	90

Можно ли утверждать, что по мере взросления студенты меньше посещают занятия?

Проверить нулевую гипотезу об отсутствии такого тренда с уровнем значимости $\alpha_0 = 0.05$ в предположении наличия фактора времени.

2. Найти с помощью эмпирической ФР $F_n^{\downarrow}(x)$, построенной по выборке $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ эмпирическую ФР случайной величины ξ^2 .
3. Найти с помощью эмпирической ФР $F_n^{\downarrow}(x)$, построенной по выборке $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ вероятность события: $P(\xi^3 < \xi)$.
4. Найти $D(I\{A\} + I\{B\})$, если события A и B независимы.
5. Пусть $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ - выборка. $f_{\xi}(x)$ - оцениваемая плотность распределения вероятностей. $\Delta_1, \Delta_2, \dots, \Delta_s$ - конечное разбиение. Найти $Mv\{\Delta_1 \cup \Delta_2 \cup \Delta_3\}$, где $v\{A\}$ - частота заполнения множества A .

6. $f_n^{\downarrow}(x)$ - гистограммная оценка плотности. Найти $\int_{-\infty}^{\infty} f_n^{\downarrow}\left(\frac{x-3}{5}\right) dx$.

7. Чему равна статистика τ Кендалла для двух наборов данных: $\vec{x} = (x_1, x_2, \dots, x_n)$ и $\vec{y} = (y_1, y_2, \dots, y_n)$, если $x_i = i$, $y_i = i^2$ для всех $i = 1, 2, \dots, n$?

8. Для данных: $\vec{x} = (1, 2, 3, 5, 7, 1)$, $\vec{y} = (0, 2, 3, 1)$ посчитать ранг $r_1^{(y)}$. Если для проверки гипотезы об отсутствии сдвига $H_0: \Delta = Mx - My = 0$ против альтернативы $H_1^-: \Delta < 0$

$$W = \sum_{i=1}^n r_i^{(x)}$$

применяется статистика W , то где должна находиться критическая область слева или справа? Ответ обоснуйте.

Критерий оценки – правильность и полнота ответа на вопросы адекватность приведенных примеров. Оценка выставляется по шкале от 0 до 30 баллов: Первая задача оценивается в 6 баллов, 2-6 задачи – 4 балла, седьмое и восьмое задания оцениваются по 2 балла каждое. Экзамен считается сданным при оценке не ниже 20 баллов.

6.2.2. Контрольная работа № 1

Контрольная работа предназначена для выявления качества усвоения практических знаний по четырем темам дисциплины, а именно: «Оценка функции распределения», «Оценка плотности распределения», «Задача о положении», «Задача о рассеянии». Эти разделы являются основополагающими для последующего освоения предмета.

Контрольная работа включает в себя 4 практические задачи, на которые студент должен дать исчерпывающий письменный ответ. Цель первой и второй задачи – оценивание закона распределения. Третья и четвертая задача тематически посвящены проверке гипотез о сдвиге и масштабе.

Перечень задач первых трех вариантов:

Вариант 1.

1. По выборке $\vec{\xi} = (10, 7, 1, 2, 5, 3, 3, 4, 5, 1)$ построить $F_n^H(x)$ и с уровнем значимости $\alpha_0 = 0.1$ проверить гипотезу согласия: $H_0: F_{\xi}(x) = 0.01x(20-x)$ при $x \in [0; 10]$.
2. По группированным данным (см. табл.2) построить гистограмму, найти визуальное распределение, наиболее подходящее под гистограмму и проверить гипотезу согласия критерием Пирсона χ^2 .
3. Проверить гипотезу $H_0: \gamma = \sigma_1^2 / \sigma_2^2 = 1$ против альтернативы $H_1^+: \gamma = \sigma_1^2 / \sigma_2^2 > 1$ по 1-ой и 2-ой строке таблицы 1 с уровнем значимости $\alpha_0 = 0.1$.
4. Проверить гипотезу $H_0: \Delta = m_1 - m_3 = 0$ против альтернативы $H_1: \Delta \neq 0$ по 1-ой и 3-ой строке таблицы 1 с уровнем значимости $\alpha_0 = 0.05$.

Таблица 1.

2	2	-4	10	-1	1	8	3	-2	5	2	2	
4	0	2	2	2	2	2	0	3	-2	4		
0	-1	2	-3	-3	-6	-1	-7	-2	0	-3	-6	7

Вариант 2.

1. По выборке $\vec{\xi} = (10, 7, 1, 8, 5, 3, 7, 4, 5, 1)$ построить $F_n^H(x)$ и с уровнем значимости $\alpha_0 = 0.05$ проверить гипотезу согласия: $H_0: F_{\xi}(x) = 0.01x^2$ при $x \in [0; 10]$.
2. По группированным данным (см. табл.2) построить гистограмму, найти визуальное распределение, наиболее подходящее под гистограмму и проверить гипотезу согласия критерием Пирсона χ^2 .
3. Проверить гипотезу $H_0: \gamma = \sigma_1^2 / \sigma_2^2 = 1$ против альтернативы $H_1^-: \gamma = \sigma_1^2 / \sigma_2^2 < 1$ по 1-ой и 2-ой строке таблицы 1 с уровнем значимости $\alpha_0 = 0.01$.
4. Проверить гипотезу $H_0: \Delta = m_1 - m_3 = 0$ против альтернативы $H_1^+: \Delta > 0$ по 1-ой и 3-ой строке таблицы 1 с уровнем значимости $\alpha_0 = 0.05$.

Таблица 1.

3	-3	1	7	-1	1	6	-6	2	-3	-1	-1
-1	1	2	2	1	-1	1	2	-2	-3	3	
0	4	-2	1	8	12	8	8	-3	-7	1	

Вариант 3.

1. По выборке $\vec{\xi} = (10, 1, 1, 8, 2, 3, 7, 4, 5, 1)$ построить $\hat{F}_n(x)$ и с уровнем значимости $\alpha_0 = 0.1$ проверить гипотезу согласия: $H_0: F_{\xi}(x) = 0.1x$ при $x \in [0; 10]$.
2. По группированным данным (см. табл.2) построить гистограмму, найти визуальное распределение, наиболее подходящее под гистограмму и проверить гипотезу согласия критерием Пирсона χ^2 .
3. Проверить гипотезу $H_0: \gamma = \sigma_1^2 / \sigma_2^2 = 1$ против альтернативы $H_1: \gamma = \sigma_1^2 / \sigma_2^2 \neq 1$ по 1-ой и 2-ой строке таблицы 1 с уровнем значимости $\alpha_0 = 0.1$.
4. Проверить гипотезу $H_0: \Delta = m_1 - m_3 = 0$ против альтернативы $H_1: \Delta < 0$ по 1-ой и 3-ой строке таблицы 1 с уровнем значимости $\alpha_0 = 0.05$.

Таблица 1.

7	5	1	0	-4	6	12	0	10	13	3	9	7
-1	1	1	3	2	1	3	-2	0	1	3	-1	
0	1	-2	-3	0	7	9	3	2	7	7	0	

Таблица 2. Эмпирические частоты.

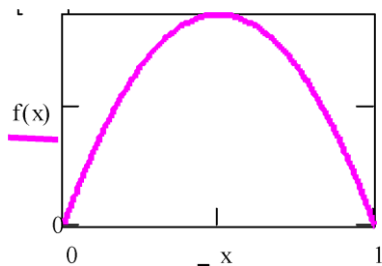
№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
[0;0.2)	70	8	65	16	2	98	21	78	11	60	71	16	9	1	64	99	61	22
[0.2;0.4)	58	25	33	48	11	60	50	21	30	55	57	49	24	12	32	59	54	49
[0.4;0.6)	40	42	8	70	30	29	60	1	45	44	41	71	41	31	8	28	43	62
[0.6;0.8)	25	56	31	50	60	10	48	20	54	29	24	50	58	61	35	12	31	50
[0.8;1]	7	69	63	16	97	3	21	80	60	12	7	14	68	95	61	2	11	17

Значения функции $\lambda_p: p = P(D_n = \sup |\hat{F}_n(x) - F(x)| > \lambda_p)$

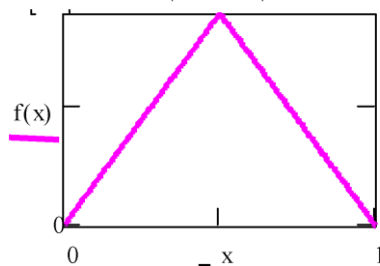
n	p			n	p		
	0,10	0,05	0,01		0,10	0,05	0,01
1	0,950	0,975	0,995	19	0,271	0,301	0,361
2	776	842	929	20	265	294	352
3	636	708	829	25	238	264	317
4	565	624	734	30	218	242	290
5	509	563	669	35	242	224	269
6	468	519	617	40	189	210	252
7	436	483	576	45	179	198	238
8	410	454	542	50	170	188	226
9	387	430	513	55	162	180	216
10	369	409	489	60	155	172	207
11	352	391	468	65	149	166	199
12	338	375	449	70	144	160	192
13	325	361	432	75	139	154	185
14	314	349	418	80	135	150	179
15	304	338	404	85	131	145	174
16	295	327	392	90	127	141	169
17	286	318	381	95	124	137	165

Рис. 1

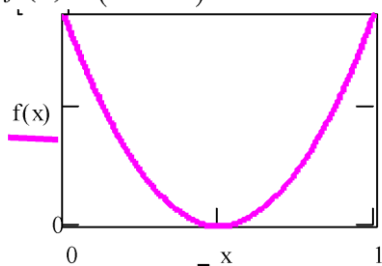
1. $f(x) \sim x(1-x)$



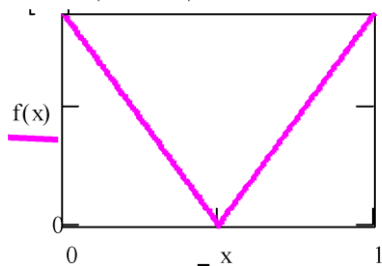
2. $f(x) \sim 0.5 - |0.5 - x|$



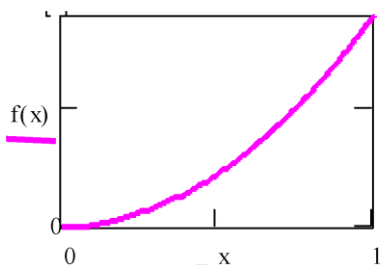
3. $f(x) \sim (x-0.5)^2$



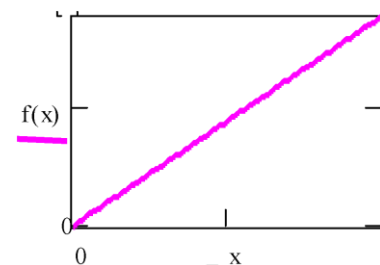
4. $f(x) \sim |x-0.5|$



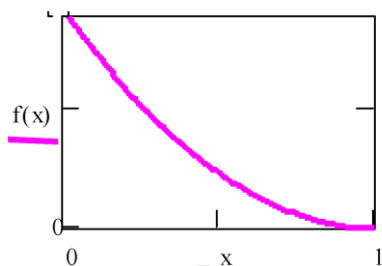
5. $f(x) \sim x^2$



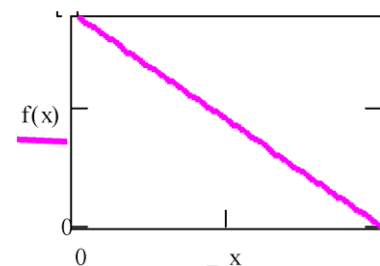
6. $f(x) \sim x$



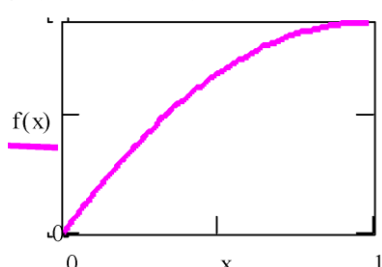
7. $f(x) \sim (1-x)^2$



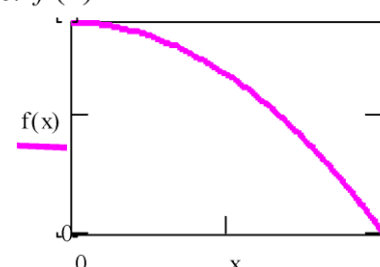
8. $f(x) \sim 1-x$



9. $f(x) \sim x(2-x)$



10. $f(x) \sim 1-x^2$



Критерий оценки – правильность и полнота ответа на задачу. Оценка выставляется по шкале от 0 до 25 баллов. Первая и вторая задачи оцениваются в 5 баллов, третья – 7, четвертая в 8 баллов соответственно. Контрольная считается написанной успешно при оценке не ниже 15 баллов.

6.2.3. Контрольная работа № 2

Контрольная работа предназначена для выявления качества усвоения практических знаний по двум темам дисциплины, а именно: «Ранги и ранговая корреляция», «Дисперсионный анализ». Эти разделы являются основополагающими для последующего освоения предмета.

Контрольная работа включает в себя 2 практические задачи, на которые студент должен дать исчерпывающий письменный ответ. Первая задача требует нахождения среднего значения функции от рангов. Тема второй задачи – одно- и двухфакторный дисперсионный анализ.

Перечень задач первых трех вариантов:

Вариант 1.

1. Найти распределение статистики $\sum_{i=1}^n \sum_{j=i}^n r_j$ при условии H_0 (случайность) при $n = 3$. Чувствительна ли статистика к упорядочиванию рангов? Найти м.о. и дисперсию статистики при H_0 $n = 3$.
2. Швейная фабрика выпускает изделия А,В,С,Д,Е тремя возможными способами. В таблице представлена информация о времени обработки изделий каждым из способов в минутах. Проверить нулевую гипотезу об отсутствии разницы в способах обработки с уровнем значимости $\alpha_0 = 0.1$

	А	В	С	Д	Е
І способ	3	5	10	4	2
ІІ способ	4	6	11	5	4
ІІІ способ	4	5	9	3	4

Вариант 2.

1. Найти распределение статистики $\sum_{i=1}^n \sum_{j=1}^i r_j$ при условии H_0 (случайность) при $n = 3$. Чувствительна ли статистика к упорядочиванию рангов? Найти м.о. и дисперсию статистики при H_0 $n = 3$.
2. Швейная фабрика выпускает изделия А,В,С,Д,Е. В таблице представлена информация о времени обработки изделий каждым из 3-х операторов швейных машин. Проверить нулевую гипотезу об отсутствии разницы во времени изготовления изделий с уровнем значимости $\alpha_0 = 0.1$ в предположении одинаковой квалификации операторов

	А	В	С	Д	Е
І оператор	3	5	10	4	2
ІІ оператор	4	6	11	5	4

Вариант 3.

1. Найти распределение статистики $\prod_{i=1}^n \sum_{j=1}^i r_j$ при условии H_0 (случайность) при $n = 3$. Чувствительна ли статистика к упорядочиванию рангов? Найти м.о. и дисперсию статистики при H_0 $n = 3$.

2. С помощью трех датчиков измеряется температура (результаты измерений представлены в таблице). Одинаково ли откалиброваны датчики или все-таки имеется систематическая ошибка хотя бы у одного из них? Проверить нулевую гипотезу об отсутствии ошибки с уровнем значимости $\alpha_0 = 0.05$.

1 датчик	98	99	95	96	97			
2 датчик	95	96	94	99	98	95	99	99
3 датчик	90	99	95	96				

Критерий оценки – правильность и полнота выполнения всех шагов решения задачи. Оценка выставляется по шкале от 0 до 25 баллов. Первая задача оценивается в 10 баллов, вторая – в 15. Контрольная считается написанной успешно при оценке не ниже 15 баллов.

6.2.4. Реферат

Реферат предназначен для выявления качества самостоятельного усвоения практических и теоретических знаний по следующим темам дисциплины: «Порядковые статистики», «Непараметрический регрессионный анализ», «Задача о положении», «Задача о рассеянии», «Дисперсионный анализ», «Другие критерии и задачи». Эти разделы являются основополагающими для освоения предмета.

Реферат включает в себя подробное изложение теоретического вопроса с изложением и обоснованием необходимых нюансов выбранной задачи и разбор одного, двух практических примеров в среде R. Реферат подразумевает выступление на выбранную тему. Кроме выступления необходим отчет о проделанной работе. Отчет должен содержать:

- Постановку задачи
- Пример практической задачи
- Исследование метода
- Пояснение на одном, двух простых примерах
- Экспериментальное обоснование применения асимптотического критерия
- Возможные оценки, связанные с критерием
- Сравнение с аналогами (по возможности)
- Решение заданной преподавателем задачи.

Темы рефератов по курсу «Дополнительные главы прикладной математики»:

2. Одновыборочная задача о положении. Критерий знаковых рангов Уилкоксона.
3. Одновыборочная задача о положении. Критерий знаков Фишера.
4. Одновыборочная задача о положении. Критерий симметрии Гупта.
5. Двухвыборочная задача о положении. Критерий ранговых сумм Уилкоксона.
6. Двухвыборочная задача о рассеянии. Критерий Ансари- Бредли.
7. Двухвыборочная задача о рассеянии. Критерий Мозеса.
8. Двухвыборочная задача о рассеянии. Критерий Миллера.
9. Однофакторный дисперсионный анализ. Критерий Краскела- Уоллиса.
10. Однофакторный дисперсионный анализ. Критерий Джонкхиера- Терпстры.
11. Двухфакторный дисперсионный анализ. Критерий Фридмана.
12. Двухфакторный дисперсионный анализ. Критерий Пейджа.
13. Двухфакторный дисперсионный анализ. Критерий Доксама.
14. Двухфакторный дисперсионный анализ. Критерий Холлендера.
15. Одна линия регрессии. Критерий и оценка Тейла.
16. Одна и две линии регрессии. Критерий параллельности и критерий Тейла.
17. Критерий двумерной симметрии Холлендера.
18. Критерий «Новое Лучше Старого».

Критерий оценки – правильность и полнота выполнения реферата. Оценка выставляется по шкале от 0 до 20 баллов. Реферат считается выполненным успешно при оценке не ниже 12 баллов.

6.3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Форма аттестации	Наименование оценочного средства	Баллы
Экзамен (100 баллов)	Контрольная работа № 1	25
	Контрольная работа № 2	25
	Реферат	20
	Ответы на экзаменационный билет	30

7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

а) основная учебная литература

1. Миносцев, В.Б. (под ред.) Курс математики для технических высших учебных заведений. Часть 4. Теория вероятностей и математическая статистика [Электронный ресурс, доступен в Электронно-библиотечной системе издательства «Лань»] : учебное пособие / В.Б. Миносцев (под ред.), Е.А. Пушкарь (под ред.), Н.А. Берков [и др.]. — Электрон. дан. — СПб. : Лань, 2013. — 304 с. — Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=32817 — Загл. с экрана.
2. Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике : учеб. пособие для бакалавров / В. Е. Гмурман. – 11-е изд., перераб. и доп. – М. : Юрайт, 2013. – 404 с. : ил. – (Бакалавр. Базовый курс) (Библиотека ИАТЭ НИЯУ МИФИ, 1 экз.)
3. Бочаров, П.П. Теория вероятностей. Математическая статистика [Электронный ресурс, доступен в Электронно-библиотечной системе издательства «Лань»] : / П.П. Бочаров, А.В. Печинкин. — Электрон. дан. — М. : Физматлит, 2005. — 296 с. — Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=59406 — Загл. с экрана.
4. Боровков А. А. Математическая статистика [Электронный ресурс, доступен в Электронно-библиотечной системе издательства «Лань»] : учебник. — Электрон. дан. — СПб. : Лань, 2010. — 704 с. — Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=3810 — Загл. с экрана.
5. Емельянов, Г.В. Задачник по теории вероятностей и математической статистике [Электронный ресурс, доступен в Электронно-библиотечной системе издательства «Лань»] : учебное пособие / Г.В. Емельянов, В.П. Скитович. — Электрон. дан. — СПб. : Лань, 2007. — 332 с. — Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=141 — Загл. с экрана.
6. Свешников, А.А. Сборник задач по теории вероятностей, математической статистике и теории случайных функций [Электронный ресурс, доступен в Электронно-библиотечной системе издательства «Лань»] : учебное пособие. — Электрон. дан. — СПб. : Лань, 2013. — 446 с. — Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=5711 — Загл. с экрана.
7. Хрущева, И.В. Основы математической статистики и теории случайных процессов [Электронный ресурс] : учебное пособие / И.В. Хрущева, В.И. Щербаков, Д.С. Леванова. — Электрон. дан. — СПб. : Лань, 2009. — 332 с. — Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=426 — Загл. с экрана.
8. Буре В. М. Теория вероятностей и математическая статистика [Электронный ресурс] : учебник / Буре В. М., Парилина Е. М. — Электрон. дан. — СПб. : Лань, 2013. — 416 с. —

Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=10249 — Загл. с экрана.

9. Кобзарь, А.И. Прикладная математическая статистика. Для инженеров и научных работников [Электронный ресурс] : учебное пособие. — Электрон. дан. — М. : Физматлит, 2012. — 813 с. — Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=59747 — Загл. с экрана.

б) дополнительная учебная литература

1. Гмурман В.Е. Теория вероятностей и математическая статистика. – М.: Высшая школа, 2003, – 479 с. Гриф МО РФ. (Библиотека ИАТЭ НИЯУ МИФИ, 50 экз)*
2. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике: Учебное пособие для студентов вузов. – М.: 2005. – 405 с. Гриф МО РФ.(Библиотека ИАТЭ НИЯУ МИФИ, 50 экз)*
3. Чепурко В.А. Методы статистического оценивания в АСУ / Под ред. С.К. Девятиловой. Учебное пособие по курсу "Статистические методы и модели в управлении". -Обнинск: ИАТЭ, 2009. -96 с. (Библиотека ИАТЭ НИЯУ МИФИ, 50 экз)*
4. Чепурко В.А. Методы проверки статистических гипотез в АСУ / Под ред. С.К. Девятиловой. Учебное пособие по курсу "Статистические методы и модели в управлении". - Обнинск: ИАТЭ, 2009. - 99 с. (Библиотека ИАТЭ НИЯУ МИФИ, 50 экз)*
5. Кибзун, А.И. Теория вероятностей и математическая статистика. Базовый курс с примерами и задачами [Электронный ресурс] : справочник / А.И. Кибзун, Е.Р. Горяинова, А.В. Наумов. — Электрон. дан. — М. : Физматлит, 2007. — 232 с. — Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=59479 — Загл. с экрана.
6. Лагутин, М.Б. Наглядная математическая статистика [Электронный ресурс] : . — Электрон. дан. — М. : Бинوم. Лаборатория знаний, 2013. — 473 с. — Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=56887 — Загл. с экрана.

8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет» (далее - сеть «Интернет»), необходимых для освоения дисциплины

1. Материалы открытой энциклопедии Wikipedia // Корневая URL: [http://ru.wikipedia.org/wiki/Математическая статистика](http://ru.wikipedia.org/wiki/Математическая_статистика)
2. Ресурсы электронно-библиотечной системы Центра информационно-библиотечного обеспечения учебно-научной деятельности НИЯУ МИФИ // URL: www.library.mephi.ru (по подписке)
3. Ресурсы научной электронной библиотеки elibrary.ru // URL: www.elibrary.ru (по подписке)
4. Ресурсы электронно-библиотечной системы издательства «Лань» // URL: www.e.lanbook.com (по подписке)
5. Ресурсы электронно-библиотечной системы образовательных и просветительских изданий // URL: www.iqlib.ru (по подписке)

9. Методические указания для обучающихся по освоению дисциплины

Вид учебного занятия	Организация деятельности студента
Лекция	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. Уделить внимание следующим понятиям: эмпирическая функция распределения, параметр сглаживания, метрика Колмогорова, Мизеса, Андерсона-Дарлингга, ядерная оценка плотности распределения, проекционная оценка

	плотности распределения, систематическая и случайная ошибки оценивания, оптимальный параметр сглаживания, базисные функции, метод локальной аппроксимации, простая и сложная гипотеза, нулевая гипотеза, альтернатива, ошибка первого и второго рода, тест, критерий, критическая область, ранг и инверсия, ранговая корреляция, параметры положения, сдвига и масштаба, фактор
Практические занятия	Проработка рабочей программы, уделяя особое внимание целям и задачам, структуре и содержанию дисциплины. Работа с конспектом лекций, просмотр рекомендуемой литературы. Изучение выбранной предметной области, включая задачи семинарских занятий и домашних работ.
Курсовая работа	Не предусмотрена
Контрольная работа	Ознакомиться с основной и дополнительной литературой, включая справочные издания, зарубежные источники, основополагающие термины. Попрактиковаться в решении задач по всем темам двух контрольных работ
Лабораторная работа	Не предусмотрена
Подготовка к зачету	При подготовке к экзамену в письменной форме ориентироваться на методические пособия, конспект лекций, рекомендуемую литературу и др.

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и информационных справочных систем (при необходимости)

1. Язык для статистической обработки данных и работы с графикой R для создания и демонстрации графиков законов распределения случайной величины и выполнения расчетов ИДЗ.
2. MS Excel для выполнения расчетов ИДЗ
3. Издательская система LaTeX для подготовки докладов, презентаций и учебного материала
4. Материалы открытой энциклопедии Wikipedia // Корневая URL: [http://ru.wikipedia.org/wiki/Математическая статистика](http://ru.wikipedia.org/wiki/Математическая_статистика)
5. Сайт проекта R: <http://www.r-project.org>
6. [Язык программирования R/Введение — Викиучебник](https://ru.wikibooks.org/wiki/Язык_программирования_R/Введение) https://ru.wikibooks.org/wiki/Язык_программирования_R/Введение
7. R — объектно-ориентированная статистическая среда: <http://ashipunov.info/shipunov/software/r/r-ru.htm>
8. [Графическая галерея R](http://web.archive.org/web/20130113002105/http://gallery.r-enthusiasts.com/) — примеры графики, генерируемой R: <http://web.archive.org/web/20130113002105/http://gallery.r-enthusiasts.com/>.

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Видеопроектор, компьютер, язык и среда R для статистической обработки данных и работы с графикой для создания и демонстрации графиков законов распределения случайной величины, выполнения контрольных работ и экзамена, MS Office Excel, издательская система LaTeX для подготовки докладов, презентаций и учебного материала.

12. Иные сведения и (или) материалы

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

Часов в интерактивной форме – 26.

Лекционный материал основан на большом количестве задач непараметрической математической статистики. Некоторые лекции сопровождаются презентациями, к примеру, разъясняющими статистические методы и модели. С привлечением языка и среды разработки R поясняются необходимые методы обработки статистической информации. Вариация входных параметров закона распределения, объема выборки и графическое сопровождение позволяет

студенту лучше понять смысл получаемых оценок. Также в среде разработки R выполняется контрольные работы, готовится реферат и экзамен.

В ходе практических занятий происходит публичное обсуждение каждой решаемой задачи и статистического метода, его достоинства и недостатки. При этом студенты высказывают свои мнения и дополняют построенную статистическую модель или предлагают свои модели, в той ситуации, когда задача этого требует.

После проведения контрольных работ на консультациях проводится разбор допущенных студентами ошибок.

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

Самостоятельно изучаются некоторые темы математической статистики при подготовке реферата. Для изучения используется приведенная в списке основная и дополнительная литература. Контроль освоения материала осуществляется в ходе экзамена.

№	Тема и часть, изучаемая (осваиваемая) самостоятельно
1.1	Оценка функции распределения. Статистики Колмогорова, Мизеса, Андерсона–Дарлинга. Теорема Гливленко-Кантелли. Построение эмпирической ф.р. в среде R имеющимися средствами.
1.2	Оценка плотности распределения. Построение гистограмм и ядерных оценок в среде R имеющимися средствами.
1.3	Порядковые статистики. Непараметрические доверительные интервалы.
2.1	Ранги и ранговые критерии в среде R. Некоторые свойства рангов, инверсий и статистик от них.
2.2	Непараметрический регрессионный анализ. Прикладные задачи идентификации. Метод локальной аппроксимации.
3.1	Задача о положении. Критерий Фишера, Вилкоксона, знаковых рангов Вилкоксона, Гупта и оценки, связанные с ними.
3.2	Задача о рассеянии. Критерий Ансари-Бредли, Мозеса, Миллера и оценки, связанные с ними.
3.3	Дисперсионный анализ. Критерий Краскела-Уоллеса, Джонкхиера-Терпстры, Фридмана и др., и оценки, связанные с ними.
3.4	Другие критерии и задачи. Биномиальный критерий, критерии Холлендера-Прошана, критерии нормальности и др. критерии и оценки, связанные с ними.

Вопросы и задания для самоконтроля по всем темам:

1. Смоделировать в среде R выборку известного распределения заданного объема.
2. Построить графики эмпирической функции распределения, гистограммы, полигона частот, ядерной оценки плотности распределения.
3. Найти числовые выборочные характеристики.
4. Вариацией входных параметров встроенных функций определить их возможности.
5. Сравнить с методикой выполнения тех же задач в других средах.
6. Разобраться с возможностями МНК в среде R.
7. Решить задачу линейного регрессионного анализа с предложенными преподавателем базисными функциями в среде R.
8. Как построить д.и. для квантили?
9. Как найти м.о. числа инверсий, суммы рангов?
10. Является ли сумма рангов с.в.? А если есть связи?
11. Чем критерий Вилкоксона отличается от критерия знаковых рангов Вилкоксона?

12. В каких задачах необходимо применять критерий Краскела-Уоллеса, а в каких – Фридмана?
13. Что такое фактор в дисперсионном анализе?
14. Означает ли равенство двух параметров положения (масштаба) однородность данных?

12.3. Краткий терминологический словарь

Выборка	Массив независимых одинаково распределенных с.в.
Объем выборки	Размер массива
Статистика	Функция от выборки
Оценка параметра	Статистика, приближенно оценивающая неизвестный параметр
Несмещенная оценка	Оценка, условное м.о. которой равно оцениваемому параметру
Состоятельная оценка	Оценка, в том или ином вероятностном смысле приближающаяся к неизвестному при росте объема выборки
Эффективная оценка	Несмещенная оценка с минимально возможной дисперсией
Асимптотически нормальная оценка	Оценка, распределение которой приближается к нормальному закону при росте объема выборки
Доверительный интервал	Интервал со случайными границами, накрывающий оцениваемый параметр с вероятностью не ниже заданной
Доверительная вероятность (надежность)	Уровень вероятности, определяющий длину д.и. Обычно берется равным – 90, 95, 100%
Статистическая гипотеза	Утверждение о законе распределения с.в., о виде зависимости и т.д.
Нулевая гипотеза	Основная гипотеза
Альтернатива	Гипотеза, хотя бы в чем- то противоречащая нулевой гипотезе