

Федеральное государственное бюджетное образовательное учреждение высшего
профессионального образования
Московский государственный университет имени М.В. Ломоносова
Факультет биоинженерии и биоинформатики

УТВЕРЖДАЮ

Декан
факультета биоинженерии
и биоинформатики,
академик

_____/В.П. Скулачев /

« ____ » _____ 20__ г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Наименование дисциплины:

Математическая статистика

Уровень высшего образования:

специалитет

Направление подготовки (специальность):

06.05.01 Биоинженерия и биоинформатика

Форма обучения:

очная

Рабочая программа рассмотрена и одобрена

Ученым советом факультета

(протокол № _____, _____)

Москва 20__

Рабочая программа дисциплины разработана в соответствии с самостоятельно установленным МГУ образовательным стандартом (ОС МГУ) для реализуемых основных профессиональных образовательных программ высшего образования по специальности 06.05.01 «Биоинженерия и биоинформатика» (программы специалитета) в редакции приказа МГУ от 30 декабря 2016 г.

Год (годы) приема на обучение – 2016, 2017, 2018, 2019.

© Факультет биоинженерии и биоинформатики МГУ имени М.В. Ломоносова

Программа не может быть использована другими подразделениями университета и другими вузами без разрешения факультета.

Цель и задачи дисциплины

Цель курса - ознакомить студентов с основными концепциями прикладной математической статистики.

Задачи курса: ожидается, что студенты будут способны читать и понимать статистические задачи, выполнять статистические тесты, выбирать статистические тесты для своих частных задач, графически представлять данные, планировать статистические исследования.

1. Место дисциплины в структуре ОПОП ВО – базовая часть, математический и естественно-научный цикл, курс II – семестр 4.

2. Входные требования для освоения дисциплины, предварительные условия (если есть):

Знание таких дисциплин, как «Теория вероятностей».

3. Планируемые результаты обучения по дисциплине:

знать основы математической статистики для решения задач анализа и интерпретации экспериментальных данных, принципы и методы статистического оценивания числовых характеристик и параметров распределений наблюдаемых случайных величин, принципы и методы проверки статистических гипотез о параметрах модели;

уметь строить статистические модели явлений и исследовать их на основе результатов измерительных экспериментов, делать статистические выводы;

владеть методами оптимального принятия статистических решений, статистического оценивания неизвестных законов распределений и навыками применения современных методов математической статистики для решения задач анализа и интерпретации экспериментальных данных

4. Формат обучения – лекционные и семинарские занятия

5. Объем дисциплины составляет 3 з.е., в том числе 64 академических часа, отведенных на контактную работу обучающихся с преподавателем, 44 академических часа на самостоятельную работу обучающихся.

6. Краткое содержание дисциплины (аннотация):

Данный курс является введением в математическую статистику как продолжение курса по теории вероятностей. В курсе рассматривается широкий спектр тем в приложении к задачам из повседневной жизни и, в особенности, к биомедицинским задачам. Материал курса предлагается в виде четырех последовательных модулей, каждый из которых состоит из лекции и практических занятий. В практикуме используется язык R-statistics и приложение R-studio. Курс направлен на изучение предмета математической статистики, а не программного обеспечения, поэтому языки программирования в нем задействованы лишь инструментально.

Наименование и краткое содержание разделов и тем дисциплины, Форма промежуточной аттестации по дисциплине	Всего (часы)	В том числе	
		Контактная работа (работа во взаимодействии с преподавателем) Виды контактной работы, часы	Самостоятельная работа обучающегося, часы (виды самостоятельной работы – эссе, реферат, контрольная работа и пр. – указываются при необходимости)

		Занятия лекционного типа	Занятия семинарского типа	Всего	
Раздел I. Описательные статистики и исследовательский анализ данных Тема 1. Исследовательский анализ данных Тема 2. Семплирование. Тема 3. Планирование эксперимента	21	6	6	12	9
Раздел II. Теория вероятностей Тема 1. Вероятность Тема 2. Предельные теоремы	18	6	6	12	6
Раздел III. Введение в проверку гипотез Тема 1. Концепция тестирования гипотез. Тема 2. Доверительные интервалы. Тема 3. Спаренные выборки.	26	8	8	16	10
Раздел IV. Проверка гипотез и непараметрические тесты Тема 1. Оценки дисперсии. Тема 2. Непараметрические тесты. Тема 3. Непараметрические тесты, продолжение. Тема 4. Оценка объема выборки. Тема 5. Поправки на множественное тестирование, ошибки в семействах гипотез.	39	12	12	24	15
Промежуточная аттестация - экзамен					4 (количество часов, отведенных на промежуточную аттестацию)
Итого	108	32	32	64	44

7. Фонд оценочных средств (ФОС) для оценивания результатов обучения по дисциплине

7.1. Типовые контрольные задания или иные материалы для проведения текущего контроля успеваемости.

Не предусмотрено

7.2. Типовые контрольные задания или иные материалы для проведения промежуточной аттестации.

1. Макдоналд и Крейтман в 1991 году секвенировали ген алкогольдегидрогеназы в нескольких особях трех видов дрозофил. Вариативные позиции классифицировались как синонимичные (последовательность аминокислот не изменялась) и несинонимичные, а также полиморфные (различаются между особями одного вида) и фиксированные. В отсутствие естественного отбора, отношение числа синонимичных замен к несинонимичным должно быть приблизительно одинаковым в полиморфных и фиксированных сайтах. В наблюдаемой ими выборке было 43 синонимичных полиморфных, 2 несинонимичных полиморфных, 17 синонимичных фиксированных и 7 несинонимичных фиксированных замен.

а. (1 балл) Чему равны ожидаемые частоты синонимичных и несинонимичных, полиморфных и фиксированных замен? Назовите тест, который нельзя использовать для этой таблицы сопряженности.

- b. (2 балла) Протестируйте на 1% уровне значимости есть ли ассоциация между синонимичностью и полиморфностью.
- c. (1 балл) Если тест Макдональда-Крейтмана применить к 100 генам, включая алкоголь дегидрогеназу, чему должно быть равно p-значение с учетом поправки на множественное тестирование? Назовите поправку на множественное тестирование.
2. Кастер и Галли на легкомоторном самолете проследили за голубыми и белыми цаплями от мест гнездования до мест кормления на озере Пелтье штат Миннесота и записали на каких поверхностях садилась каждая птица. Достаточно ли оснований считать на 5% уровне значимости, что два вида птиц различаются по своим предпочтениям?
- Да, так как $p\text{-значение} < 0.01$
 - Нет, так как $p\text{-значение} < 0.01$
 - Да, так как $p\text{-значение} < 0.02$
 - Нет, так как $p\text{-значение} < 0.02$
 - Ни одно из перечисленного

Шкала и критерии оценивания результатов обучения по дисциплине.

Результаты обучения	«Неудовлетворительно»	«Удовлетворительно»	«Хорошо»	«Отлично»
Знания: основ математической статистики для решения задач анализа и интерпретации экспериментальных данных, принципов и методов статистического оценивания числовых характеристик и параметров распределений наблюдаемых случайных величин, принципов и методов проверки статистических гипотез о параметрах модели	Знания отсутствуют	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
Умения: строить статистические модели явлений и исследовать их на основе результатов измерительных экспериментов, делать статистические выводы	Умения отсутствуют	В целом успешное, но не систематическое умение	В целом успешное, но содержащее отдельные пробелы умение (допускает неточности не принципиального характера)	Успешное и систематическое умение
Владения:	Навыки владения	Наличие	В целом,	Сформированные

методами оптимального принятия статистических решений, статистического оценивания неизвестных законов распределений и навыками применения современных методов математической статистики для решения задач анализа и интерпретации экспериментальных данных	отсутствуют	отдельных навыков (наличие фрагментарного опыта)	сформированные навыки (владения), но используемые не в активной форме	ые навыки (владения), применяемые при решении задач
--	-------------	--	---	---

8. Ресурсное обеспечение:

- Перечень основной и дополнительной литературы

1	Samuels, Witmer, and Schaffner	Statistics for Life sciences	California Polytechnic State University	Fourth Edition	2006
2	Гмурман В.Е.	Руководство к решению задач по теории вероятностей и математической статистике	Высшая школа		1988
3	Peck, Olsen Devore	Introduction to Statistics and Data Analysis	Thomson Higher Education	Third Edition	2008

- Перечень лицензионного программного обеспечения (при необходимости)
- Перечень профессиональных баз данных и информационных справочных систем
- Перечень ресурсов информационно-телекоммуникационной сети «Интернет» (при необходимости)
- Описание материально-технического обеспечения.
Лекционные и семинарские аудитории
Проектор, доска, персональные компьютеры
Статистические таблицы, калькуляторы