

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

**Федеральное государственное бюджетное учреждение науки
Институт общей генетики им. Н.И. Вавилова
Российской академии наук
(ИОГен РАН)**

«ПРИНЯТО»

«УТВЕРЖДАЮ»

На заседании Ученого совета ИОГен РАН

Директор ИОГен РАН

Протокол № 1 от «19» февраля 2019 г. д.б.н.  А.М. Кудрявцев



**РАБОЧАЯ ПРОГРАММА
ПО ДИСЦИПЛИНЕ
Б1.В.ОД2
«БИОСТАТИСТИКА»**

Направление подготовки: 06.06.01 Биологические науки

Уровень образования: высшее образование – подготовка кадров высшей
квалификации

Квалификация выпускника: Исследователь. Преподаватель-исследователь.

Разработчик:

д.б.н. А.В. Рубанович

Москва, 2019 г.

Рабочая программа составлена на основании федерального государственного образовательного стандарта, разработанного для реализации основных профессиональных образовательных программ высшего образования - программ подготовки научно-педагогических кадров в аспирантуре по направлению подготовки кадров высшей квалификации 06.06.01 «Биологические науки».

Согласно Федеральному государственному образовательному стандарту высшего образования по направлению подготовки 06.06.01 Биологические науки (уровень подготовки кадров высшей квалификации), утвержденному приказом Минобрнауки РФ № 871 от 30 июля 2014 г., и учебному плану аспирантов, разработанного на основе этого стандарта, дисциплина «Биостатистика» является второй обязательной учебной дисциплиной модуля вариативной части Блока 1 образовательной программы по направленности (профилю) 03.02.07 Генетика.

Объём курса составляет 5 зачетных единиц или 180 академических часа, из них 40 академических часов лекций, 40 практических занятий (семинаров), 100 академических часов самостоятельной внеаудиторной работы аспирантов, включая подготовку к зачету.

1. Цели и задачи

Цель дисциплины:

освоение специфических методов статистической обработки данных в области биологии, медицины и сельского хозяйства.

Задачи дисциплины:

- формирование базовых знаний о методах статистического анализа результатов биологических экспериментов, обработка больших массивов современных генетических данных, оценка генетических рисков и показателей наследуемости, статистика биомаркеров, ROC-анализ, анализ множественности тестов, мета-анализ, построение байесовских оценок;
- практическое освоение аспирантами компьютерных методов проведения статистического анализа;
- формирование у аспирантов основных навыков статистического анализа и приобретение ими практического опыта, необходимого для проведения самостоятельных научных исследований в области генетики, эпидемиологии и физико-химической биологии.

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

Дисциплина «Биостатистика» включает в себя разделы, которые могут быть отнесены к вариативным частям Учебного плана Обязательных дисциплин - ОД1, ОД2, ОД3 и дисциплин по выбору – ВД1, ВД2 и базируется на дисциплинах Модулей 2 и 3 вариативной части.

3. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы.

Освоение дисциплины направлено на формирование следующих универсальных (УК), общепрофессиональных (ОПК) и профессиональных компетенций (ПК) аспиранта:

Универсальные компетенции:

- Способность к критическому анализу и оценке современных научных достижений, генерированию новых идей при решении исследовательских и практических задач, в том числе в междисциплинарных областях (УК-1);
- Способность проектировать и осуществлять комплексные исследования, в том числе междисциплинарные, на основе целостного

системного научного мировоззрения с использованием знаний в области истории и философии науки (УК-2);

- Готовность участвовать в работе российских и международных исследовательских коллективов по решению научных и научно-образовательных задач (УК-3);

- Готовность использовать современные методы и технологии научной коммуникации на государственном и иностранном языках (УК-4);

- Способность планировать и решать задачи собственного профессионального и личностного развития (УК-5);

Общепрофессиональные компетенции:

- способность самостоятельно осуществлять научно-исследовательскую деятельность в соответствующей профессиональной области с использованием современных методов исследования и информационно-коммуникационных технологий (ОПК-1);

- готовность к преподавательской деятельности по основным образовательным программам высшего образования (ОПК-2).

Профессиональные компетенции:

- Способность к самостоятельному проведению научно-исследовательской работы и получению научных результатов, удовлетворяющих установленным требованиям к содержанию диссертаций на соискание ученой степени кандидата наук по направленности 03.02.07 Генетика (ПК-1);

- Обладание представлениями о фундаментальных основах биологических процессов, форм и методов научного познания (ПК-2);

- Способность приобретать новые знания с использованием современных научных методов и владение ими на уровне, необходимом для решения задач, возникающих при выполнении профессиональных функций (ПК-3);

- Обладание опытом профессионального участия в научных дискуссиях, умение представлять полученные в исследованиях результаты в виде отчетов и научных публикаций (стендовые доклады, рефераты и статьи в периодической научной печати) (ПК-4);

- Владение методами отбора материала, преподавания и основами управления процессом обучения фундаментальной биологии (ПК-5).

В результате освоения дисциплины, обучающиеся должны:

знать:

- основные методы оценки статистической значимости;
- методы учета множественности сравнений;
- методы мета-анализа;
- статистические характеристики ассоциативных тестов;
- ROC-анализ;
- методы оценки наследуемости и генетических рисков;
- методы сокращения числа переменных при анализе больших массивов данных;
- методы классификации данных;
- основы байесовского анализа данных.

уметь:

- пользоваться Интернетом и справочной литературой по биостатистике научного и прикладного характера для быстрого поиска необходимых данных и понятий;
- сравнивать между собой методы статистической обработки и адекватно оценивать их применимость;
- применять основные методы биостатистики в научных исследованиях;
- применять основные методы биостатистики при работе в лаборатории.

владеть:

- навыками обработки больших массивов данных;
- культурой компьютерного анализа статистической значимости результатов генетических и медико-биологических экспериментов

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

4.1. Объем дисциплины и виды учебной работы:

Всего часов	Аудиторные занятия (час):	Практические занятия (семинары)	Самостоятельная работа	Зачеты
180	40	40	98	2

4.2. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№ п/п	Наименование тем	Аудиторные занятия (час)	Практические занятия (семинары)	Самост. работа
1	Структура биологических данных и описательные статистики	15	15	30
2	Анализ сопряженности признаков	6	6	20
3	Многомерные методы статистического анализа	15	15	30
4	Байесовская статистика	4	4	18
	Всего	40	40	98
	Зачеты			2
	Итого			180

4.3. Содержание курса

Тема 1. Структура биологических данных и описательные статистики

- Организация файлов и управление данными в программах EXCEL, SPSS и STATISTICA. Описательные статистики.
- Некоторые приемы быстрых статистических вычислений.
- Проверка статистических гипотез. Точные и опосредованные критерии. Ошибки I и II рода. Мощность теста. Множественные сравнения. Контроль ошибок I рода.
- Страты и парадокс Симпсона.
- Параметрические и непараметрические критерии сравнения. Дисперсионный анализ.

Тема 2. Анализ сопряженности признаков

- Регрессионный анализ. Анализ остатков. Частные корреляции и конфаундеры.
- Сопряженность качественных признаков. Отношение шансов и относительный риск. Статистика биомаркеров.
- Оценки чувствительности и специфичности теста. ROC-анализ.

Тема 3. Многомерные методы статистического анализа

- Множественный регрессионный анализ.
- Методы сокращения числа предикторов.
- Парадокс Фридмана.
- Оценки наследуемости и генетического риска.
- Проблема «missing heritability». Факторный анализ.
- Метод главных компонент.
- Методы классификации. Кластерный анализ. Дискриминантный анализ.

Тема 4. Байесовская статистика

- Ограниченность концепции p-value. Анализ воспроизводимости результатов экспериментов. Байесовский фактор.
- Приоры. Статистика в эпидемиологии. Анализ больших выборок. Байесовские оценки частот редких событий

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Оборудование, необходимое для лекций и семинаров: компьютер Windows, MS Office и мультимедийное оборудование (проектор, звуковая система)

6. Перечень основной и дополнительной литературы, необходимой для освоения дисциплины (модуля)

Основная литература

1. Biswas, A., Datta, S., Fine, J. P., Segal, M. R. (eds.) Statistical Advances in the Biomedical Sciences Clinical Trials, Epidemiology, Survival Analysis, and Bioinformatics, 2008, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany.
2. Statistical Human Genetics. Edited by Robert C. Elston. Springer Science+Business Media, LLC 2012. <https://link.springer.com/book/10.1007/978-1-61779-555-8>

Дополнительная литература

1. Topics in Biostatistics. Edited by Walter T. Ambrosius 2007 Humana Press Inc. 999 Riverview Drive, Suite 208, Totowa, New Jersey 07512;

2. Agostino Di Ciaccio, Mauro Coli, Jose Miguel Angulo Ibañez. Advanced Statistical Methods for the Analysis of Large Data-Sets. Springer-Verlag Berlin Heidelberg, 2012 <https://link.springer.com/book/10.1007/978-3-642-21037-2>

3. Analysis of Large Data-Sets. Springer-Verlag Berlin Heidelberg, 2012

7. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

-

8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины

Научно-библиографические и патентные базы данных в области биологии, доступные по сети Интернет в бесплатном режиме - Science Citation Index (Web of Science), Medline (PubMed), Научная электронная библиотека (НЭБ), Российская патентная БД ФГУ ФИПС и американская патентная БД USPAFULL; электронные адреса крупных научных издательств, предоставляющих доступ к полным текстам текущих и архивным выпускам этих журналов.

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

Доступ в Интернет, дополнительно программное обеспечение не требуется.

10. Методические указания для обучающихся по освоению дисциплины

Аспирант, изучающий дисциплину, должен с одной стороны, овладеть общим понятийным аппаратом, а с другой стороны, должен научиться применять теоретические знания на практике. В результате изучения дисциплины аспирант должен знать основные определения дисциплины, уметь применять полученные знания для решения различных задач.

Успешное освоение курса требует:

- посещения всех занятий, предусмотренных учебным планом по дисциплине;
- ведения конспекта занятий;
- напряжённой самостоятельной работы.

Самостоятельная работа включает в себя:

- чтение рекомендованной литературы;

- проработку учебного материала, подготовку ответов на вопросы, предназначенных для самостоятельного изучения;
- решение задач, предлагаемых на занятиях;
- подготовку к выполнению заданий текущей и промежуточной аттестации.

Показателем владения материалом служит умение без конспекта отвечать на вопросы по темам дисциплины.

Важно добиться понимания изучаемого материала, а не механического его запоминания. При затруднении изучения отдельных тем, вопросов, следует обращаться за консультациями к преподавателю.

Возможен промежуточный контроль знаний аспирантов в виде решения задач в соответствии с тематикой занятий.

11. Фонд оценочных средств для проведения промежуточной аттестации по итогам обучения

Приложении А.

12. Составители программы:

д.б.н. Рубанович А.В.

Фонд оценочных средств для проведения промежуточной аттестации по итогам обучения

1. Компетенции, формируемые в процессе изучения дисциплины

Освоение дисциплины направлено на формирование у обучающегося следующих универсальных (УК), общепрофессиональных (ОПК) и профессиональных (ПК) компетенций:

Универсальные компетенции:

- Способность к критическому анализу и оценке современных научных достижений, генерированию новых идей при решении исследовательских и практических задач, в том числе в междисциплинарных областях (УК-1);

- Способность проектировать и осуществлять комплексные исследования, в том числе междисциплинарные, на основе целостного системного научного мировоззрения с использованием знаний в области истории и философии науки (УК-2);

- Готовность участвовать в работе российских и международных исследовательских коллективов по решению научных и научно-образовательных задач (УК-3);

- Готовность использовать современные методы и технологии научной коммуникации на государственном и иностранном языках (УК-4);

- Способность планировать и решать задачи собственного профессионального и личностного развития (УК-5).

Общепрофессиональные компетенции:

- способность самостоятельно осуществлять научно-исследовательскую деятельность в соответствующей профессиональной области с использованием современных методов исследования и информационно-коммуникационных технологий (ОПК-1);

- готовность к преподавательской деятельности по основным образовательным программам высшего образования (ОПК-2).

Профессиональные компетенции:

- Способность к самостоятельному проведению научно-исследовательской работы и получению научных результатов, удовлетворяющих установленным требованиям к содержанию диссертаций на

соискание ученой степени кандидата наук по направленности 03.02.07 Генетика (ПК-1);

- Обладание представлениями о фундаментальных основах биологических процессов, форм и методов научного познания (ПК-2);

- Способность приобретать новые знания с использованием современных научных методов и владение ими на уровне, необходимом для решения задач, возникающих при выполнении профессиональных функций (ПК-3);

- Обладание опытом профессионального участия в научных дискуссиях, умение представлять полученные в исследованиях результаты в виде отчетов и научных публикаций (стендовые доклады, рефераты и статьи в периодической научной печати) (ПК-4);

- Владение методами отбора материала, преподавания и основами управления процессом обучения фундаментальной биологии (ПК-5).

2. Показатели оценивания компетенций

В результате изучения дисциплины «Биостатистика» обучающийся должен:

знать:

- основные методы оценки статистической значимости;
- методы учета множественности сравнений;
- методы мета-анализа;
- статистические характеристики ассоциативных тестов;
- ROC-анализ;
- методы оценки наследуемости и генетических рисков;
- методы сокращения числа переменных при анализе больших массивов данных;
- методы классификации данных;
- основы байесовского анализа данных.

уметь:

- пользоваться Интернетом и справочной литературой по биостатистике научного и прикладного характера для быстрого поиска необходимых данных и понятий;
- сравнивать между собой методы статистической обработки и адекватно оценивать их применимость;

- применять основные методы биостатистики в научных исследованиях;
- применять основные методы биостатистики при работе в лаборатории.

владеть:

- навыками обработки больших массивов данных;
- культурой компьютерного анализа статистической значимости результатов генетических и медико-биологических экспериментов

3. Перечень типовых контрольных заданий, используемых для оценки знаний, умений, навыков

Промежуточная аттестация по дисциплине «Биостатистика» осуществляется в форме зачета. Зачет проводится устной форме.

1. Мощность статистических тестов.
2. Точный критерий Фишера.
3. FDR-метод учета множественности сравнений.
4. Реализация перестановочных статистических тестов.
5. Фиксированные и случайные модели мета-анализа.
6. Методы оценки публикационного сдвига. Графики-воронки.
7. Оценки чувствительности и специфичности маркера
8. Дизайн ассоциативных генетических исследований.
9. Отношение шансов и относительный риск.
10. Проверка значимости коэффициент детерминации.
11. Парадокс Фридмана.
12. Оценка генетического риска.
13. Оценки наследуемости по результатам GWAS.
14. Вычисление фактора Байеса.
15. Байесовские оценки частот редких событий.

Тестирование:

1. Сколько параметров у нормального распределения
 - а) 2
 - б) 3
 - в) 5
2. При каком минимальном количестве тестов необходимо делать поправку на множественное сравнение
 - а) 1
 - б) 2
 - в) 10
3. Из приведенных ниже выберите непараметрический статистический критерий
 - а) t-критерий Стьюдента
 - б) U-критерий Манна-Уитни
 - в) F- критерий Фишера

4. Критерии оценивания

Оценка отлично (5 баллов) - выставляется аспиранту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины, проявляющему интерес к данной предметной области, продемонстрировавшему умение уверенно и творчески применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка хорошо (4 балла) - выставляется аспиранту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но недостаточно грамотно обосновывает полученные результаты.

Оценка удовлетворительно (3 балла) - выставляется аспиранту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он освоил основные разделы учебной программы, необходимые для дальнейшего обучения, и может применять полученные знания по образцу в стандартной ситуации.

Оценка неудовлетворительно (2 балла) - выставляется аспиранту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных принципов и не умеет использовать полученные знания при решении типовых задач.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности.

При проведении устного зачета обучающемуся предоставляется 30 минут на подготовку. Опрос обучающегося по билету не должен превышать одного астрономического часа.