**Программа утверждена на заседании кафедры теории вероятностей**

**Протокол № 6 от 18 ноября 2015 г.**

**Рабочая программа дисциплины (модуля)**

1. Код и наименование дисциплины (модуля): Избранные главы математической статистики.

2. Уровень высшего образования – специалитет.

3. Направление подготовки: 01.05.01 Фундаментальные математика и механика. Специализация: Фундаментальная математика.

4. Место дисциплины (модуля) в структуре ООП: вариативная часть ООП. Является специальной дисциплиной (спецкурсом) для студентов 3-6 годов обучения, специализирующихся в данной научной области или смежной научной области, спецкурсом по выбору студента.

Освоение дисциплины необходимо для последующего изучения дисциплин образовательной программы: курсовая работа, научно-исследовательская практика, преддипломная практика, выпускная квалификационная работа.

5. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями выпускников)

6. Объем дисциплины (модуля) в зачетных единицах с указанием количества академических или астрономических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся:

Объем дисциплины (модуля) составляет 3 зачетных единицы, всего 108 часа, из которых 44 (46\*) часа составляет контактная работа студента с преподавателем (34 (36\*) часа занятия лекционного типа, 12 часов мероприятия текущего контроля успеваемости и промежуточной аттестации), 64 (62\*) часа составляет самостоятельная работа студента.

*\* - если специальный курс читается в нечетном семестре (продолжительность нечетного семестра 18 недель, четного семестра 17 недель).*

7. Входные требования для освоения дисциплины (модуля), предварительные условия.

Для того чтобы изучение дисциплины было возможно, обучающийся должен

1. освоить следующие дисциплины образовательной программы: математический анализ, линейную алгебру и геометрию, теорию вероятностей, математическую статистику, теорию случайных процессов, комплексный анализ, функциональный анализ.
2. обладать следующими компетенциями:

Знать: основные направления, проблемы, теории и методы современной математики.

Уметь: решать стандартные задачи математического анализа, линейной алгебры и геометрии, теории вероятностей, математической статистики, теории случайных процессов, комплексного анализа, функционального анализа, и применять идеи, использованные в их решениях, для решения аналогичных задач.

Владеть: основными понятиями и теоремами из этих разделов математики.

8. Формат обучения.

Очная форма обучения, лекционные занятия.

9. Содержание дисциплины (модуля), структурированное по темам (Перечень тем см. Приложения).

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Наименование и краткое содержание разделов и тем дисциплины (модуля),**  **форма промежуточной аттестации по дисциплине (модулю)** | **Всего**  **(часы**) | В том числе | | | | | | | | |
| **Контактная работа (работа во взаимодействии с преподавателем), часы**  из них | | | | | | **Самостоятельная работа обучающегося, часы**  из них | | |
| Занятия лекционного типа | Занятия семинарского типа | Групповые консультации | Индивидуальные консультации | Учебные занятия, направленные на проведение текущего контроля успеваемости, промежуточной аттестации | **Всего** | Выполнение домашних заданий | Подготовка рефератовит.п.. | **Всего** |
| Тема 1 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 2 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 3 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 4 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 5 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 6 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 7 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 8 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Текущий контроль успеваемости | 6 |  |  |  |  | 2 | 2 | 4 |  | 4 |
| Тема 9 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 10 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 11 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 12 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 13 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 14 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 15 | 6 | 2 |  |  |  |  | 2 | 4 |  | 4 |
| Тема 16 | 4 |  |  |  |  |  | 0 | 4 |  | 4 |
| Тема 17\* | 2\* |  |  |  |  |  |  | 2\* |  | 2\* |
| Промежуточная аттестация  *экзамен*  *зачет* | 8 (6\*) |  |  |  |  | 2 | 2 | 6(4\*) |  | 6 (4\*) |
| **Итого** | 108 | 30 |  |  |  | 4 | 34 | 74 |  | 74 |

10. Перечень учебно-методического обеспечения для самостоятельной работы студентов по дисциплине (модулю):

Конспекты лекций, списки задач к лекциям, основная и дополнительная учебная литература.

11. Фонд оценочных средств для промежуточной аттестации по дисциплине (модулю).

* Перечень компетенций:
* Описание шкал оценивания*:*

*экзамен с оценкой по пятибалльной шкале*

*зачет («зачтено» или «не зачтено»)*

* Критерии и процедуры оценивания результатов обучения по дисциплине (модулю), характеризующих этапы формирования компетенций.
* Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения, характеризующих этапы формирования компетенций. См. Приложения.

12. Ресурсное обеспечение:

Перечень основной учебной литературы: см. Приложение

Перечень дополнительной учебной литературы: см. Приложения

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»: см. Приложения.

Описание материально-технической базы: аудитории для проведения лекционных занятий.

13. Язык преподавания: русский (при необходимости – английский).

ПРИЛОЖЕНИЕ

1. ИЗБРАННЫЕ ГЛАВЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ. «Статистические методы анализа данных больших размерностей».
2. Лектор – проф. А.В. Булинский.
3. Аннотация курса: специальный курс для студентов 4-6 курсов и аспирантов. Цель лекций – ознакомить слушателей с современными методами анализа данных больших размерностей, а также рассмотреть разнообразные приложения излагаемой теории.
4. Тематическое содержание курса:

|  |  |
| --- | --- |
| Тема 1 | Методы выбора стохастической модели. |
| Тема 2 | Регрессионные модели. Метод наименьших квадратов и его обобщения. |
| Тема 3 | Метод главных компонент и его развитие. |
| Тема 4 | Метод LASSO. |
| Тема 5 | Обобщения метода LASSO. |
| Тема 6 | Логистическая регрессия. |
| Тема 7 | Метод опорных векторов и его обобщения. |
| Тема 8 | Модели, основанные на графах. Случайные деревья, леса. |
| Тема 9 | Логическая регрессия. |
| Тема 10 | Метод SCAD. |
| Тема 11 | Методы классификации данных. |
| Тема 12 | Стохастическая оптимизация. |
| Тема 13 | Кросс-валидация и бутстрэп. |
| Тема 14 | MDR метод и его обобщения. |
| Тема 15 | Перестановочные тесты. |
| Тема 16 | Проблемы устойчивости выбора значимых переменных. |
| Тема 17\* | Некоторые приложения теории в медицине и биологии. |

*\* - если специальный курс читается в нечетном семестре (продолжительность нечетного семестра 18 недель, четного семестра 17 недель).*

1. Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения, характеризующих этапы формирования компетенций.

*Программа экзамена (или вопросы к устному зачету)*

1.Методы выбора стохастической модели.

2. Регрессионные модели. Метод наименьших квадратов и его обобщения.

3. Метод главных компонент и его развитие.

4. Метод LASSO.

5. Обобщения метода LASSO.

6. Логистическая регрессия.

7. Метод опорных векторов и его обобщения.

8. Модели, основанные на графах. Случайные деревья, леса.

9. Логическая регрессия.

10. Метод SCAD.

11. Методы классификации данных и прогноза отклика.

12. Стохастическая оптимизация.

13. Кросс-валидация и бутстрэп.

14. MDR метод и его обобщения.

15. Перестановочные тесты.

16. Проблемы устойчивости выбора значимых переменных.

17. Некоторые приложения теории в медицине и биологии.

*Экзаменационные билеты формируются в виде двух вопросов из указанного списка и одной задачи.*

Образцы билетов.

**Билет №1**

1. Методы выбора стохастической модели.

2. MDR метод и его обобщения.

**Задача.** Пусть имеются элементы {1,…,10} и система подмножеств s1={1,2,5,8}, s2={2,4,5}, s3={1,2,6,8,9}. Найти коэффициент взвешенной состоятельности этой системы, предложенный П.Сомолем и Я.Нововичовой.

**Билет №2**

1**.** Регрессионные модели. Метод наименьших квадратов и его обобщения.

2. Проблемы устойчивости выбора значимых переменных.

**Задача.** Пусть отклик Y принимает значения 1 и -1, а факторы X1,…,Xn, от которых он зависит, принимают значения

в некотором конечном множестве T. Пусть дана штрафная функция ψ, отображающая множество {-1,1} в R+ . Найти все функции

f: Tn → {-1,1}, которые обеспечивают минимум ошибки Err(f):=E|Y-f(X1,…,Xn)|ψ(Y).

**Билет №3**

1**.** Метод главных компонент и его развитие.

2. Кросс-валидация и бутстрэп.

**Задача.** Доказать вариантусиленного закона больших чисел для массивов ограниченных случайных величин, которые независимы

в каждой строке этого массива.

1. Перечень основной и дополнительной учебной литературы, ресурсов информационно-телекоммуникационной сети «Интернет»:

Перечень учебной литературы:

[1] H.Abdi et al. (Eds). New Perspectives in Partial Least Squares and Related Methods. Springer, 2013.

[2] R.Arboretti, L.Corain, D.Mazzaro, L.Salmaso. Permutation Testing for Isotonic Inference on Association Studies in Genetics. Springer, 2011.

[3] V. Bolon-Canedo, N.Sanchez-Marono, A.Alonso-Betanzos. Feature Selection for High-Dimensional Data. Springer, 2015.

[4] P.Buhlemann, S. van de Geer. Statistics for High-Dimensional Data. Methods, Theory and Applications. Springer, 2011.

[5] A.Bulinski. Some statistical methods in genetics. Lecture Notes in Mathematics, v. 2120, p. 293-320, Springer, 2015.

[6] C.Giraud. Introduction to High-Dimensional Statistics. CRC, 2015.

[7] T.Hastie, R.Tibshirani, M.Wainwright. Statistical Learning with Sparsity. The LASSO and Generalizations. CRC, 2015.

[8] G.James, D.Witten, T.Hastie, R.Tibshirani. An Introduction to Statistical Learning with Applications in R. Springer, 2014.

[9] U.Stanczyk, L.C.Jane (Eds.) Feature Selection for Data and Pattern Recognition. Springer, 2015.

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

http://lib.mexmat.ru/

<http://elibrary.ru/>

<http://www.mathnet.ru/>

<http://www.sciencedirect.com/>

<http://www.ams.org/mathscinet/>

http://new.math.msu.su/department/probab/index-k.html