

Программа учебной дисциплины

Название дисциплины	Статистический анализ данных космического эксперимента
Где проводится	кафедра физики космоса НИУ ВШЭ
Автор программы	Петрукович Анатолий Алексеевич, д-р физ.-мат. наук, член-корреспондент РАН, профессор кафедры физики космоса НИУ ВШЭ, директор ФГБУН ИКИ РАН
Курс	4-ый курс бакалавриата
Модули	2-ый и 3-ий модули
Объём курса	2 часа лекция и 2 часа семинар в неделю
Элементы контроля	Экзамен

1. Аннотация дисциплины

Целью курса является формирование компетенций применения статистических методов в космических науках в части закрепления базовых знаний в области математической статистики, обучения методам статистического анализа данных, характерным для космического эксперимента, ознакомления с возможными проблемами применения стандартных методов, общими принципами построения и валидации статистических методов. Примерный перечень тем курса: история теории вероятности и математической статистики; базовые определения и понятия теории вероятности и математической статистики; методы проверки гипотез; непараметрические методы; анализ связей между случайными величинами, корреляция и регрессия; метод максимального правдоподобия и байесовский подход; временные ряды и спектральный анализ; АРСС методы и цифровые фильтры; регрессия и классификация сложных данных.

2. Программа дисциплины

Цель дисциплины:

Формирование компетенций применения статистических методов в космических науках.

Задачи дисциплины:

- Формирование (восстановление) базовых знаний в области математической статистики.
- Обучение методам статистического анализа данных, характерным для космического эксперимента.
- Ознакомление с общими принципами построения и валидации статистических методов.

Дисциплина «Статистический анализ данных космического эксперимента» охватывает следующие темы:

- история теории вероятности и математической статистики;
- базовые понятия теории вероятности;
- базовые понятия математической статистики;
- проверка гипотез;
- непараметрические методы;
- анализ связей между случайными величинами;
- метод максимального правдоподобия и байесовский подход;
- временные ряды и спектральный анализ;
- АРСС методы;
- цифровые фильтры;
- регрессия и классификация сложных данных.

3. Элементы контроля и правила оценивания

Оценивание усвоенных студентами знаний и навыков по теме курса состоит из двух промежуточных контрольных и одного экзамена. Итоговая оценка студента за курс складывается из оценок за контрольные и экзамен в соотношении 3:3:4.

Промежуточные контрольные нацелены в первую очередь на проверку практических навыков, полученных в ходе семинарских занятий. Задания данных контрольных работ являются вариациями разбираемых на семинарах задач по темам курса. В ходе решения задач проверяются знания и навыки студентов по практической реализации методов статистического анализа данных космических экспериментов с помощью доступных программных решений на персональных компьютерах. Итогом решения задачи должен являться программный код, корректно выполняющий поставленную задачу. При этом студент обязан полностью понимать представленный программный и быть способен объяснить решение задачи.

В ходе решения задач контрольной работы допускается использование любых доступных справочных материалов. Оценка за складывается из правильности и корректности решенной задачи а также из способности студента объяснить представленное решение.

Экзамен проводится в устной форме по билетам. В каждом билете представлен один теоретический вопрос.

Во время подготовки ответа обучающиеся могут пользоваться программой дисциплины, материалами лекций, а также любой справочной литературой и вычислительной техникой.

После ответа на основной вопрос могут быть заданы дополнительные вопросы для проверки знания базовых понятий математической статистики. При необходимости для подготовки ответа может быть предоставлено дополнительное время.

Оценка отлично 10 баллов - выставляется студенту, 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 отлично 9 баллов - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка отлично 8 баллов - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, правильное обоснование принятых решений, с некоторыми недочетами.

Оценка хорошо 7 баллов - выставляется студенту, если он твердо знает материал, грамотно и, по существу, излагает его, умеет применять полученные знания на практике, но недостаточно грамотно обосновывает полученные результаты.

Оценка хорошо 6 баллов - выставляется студенту, если он твердо знает материал, грамотно и, по существу, излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности.

Оценка хорошо 5 баллов - выставляется студенту, если он в основном знает материал, грамотно и, по существу, излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач достаточно большое количество неточностей.

Оценка удовлетворительно 4 балла - выставляется студенту, 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 удовлетворительно 3 балла - выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, допускающему ошибки в формулировках базовых понятий, нарушения логической последовательности в изложении программного материала, слабо владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и с трудом применяет полученные знания даже в стандартной ситуации.

Оценка неудовлетворительно 2 балла - выставляется студенту, 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.

Оценка неудовлетворительно 1 балл - выставляется студенту, который не знает основного содержания учебной программы дисциплины, допускает грубейшие ошибки в формулировках базовых понятий дисциплины и вообще не имеет навыков решения типовых практических задач.

4. Примеры заданий элементов контроля

Примерный перечень тематик и видов задач в промежуточных контрольных:

1. Построение основных типов распределений. Квантили.
2. Точечные и интервальные оценки параметров.
3. Проверка выборки на нормальность.
4. Анализ таблиц сопряженности.

5. Непараметрические методы (задача о сдвиге).
6. Задачи на коэффициенты корреляции и ковариации. Матрица ковариации. Ранговый коэффициент корреляции.
7. Линейная регрессия. Доверительные интервалы для параметров. Проверка значимости параметров/модели.
8. Определить выборочные характеристики стационарной случайной последовательности: математическое ожидание, дисперсия, нормированное среднее абсолютное отклонение, коэффициент асимметрии, коэффициент эксцесса.
9. Реализовать метод MVA для заданного ряда.
10. Проверка гипотезы χ^2 о нормальности временного ряда.
11. Построить периодограмму используя различные оконные преобразования, объяснить влияния вида оконного преобразования.
12. Построить временной ряд по модели АРСС.
13. Построить периодограмму на нерегулярной сетке.
14. К заданному временному ряду применить фильтрацию частот с помощью дискретного фильтра с конечной или бесконечной импульсной характеристикой.
15. Оценить параметры процесса АРСС для заданного временного ряда.

Примерный перечень контрольных вопросов в экзаменационных билетах:

1. интервальные оценки среднего (для известной и неизвестной дисперсии).
2. состоятельность, смещение, эффективность оценок на примере дисперсии
3. регрессия. Линейная регрессия для гауссовых данных (вывод формулы)
4. Корреляция. Доверительный интервал. Двухпараметрические нормальные распределения.
5. непараметрические методы. Понятие и простейшие примеры
6. Проверка гипотез о среднем. Ошибки первого и второго рода.
7. Таблицы 2x2. Критерии Пирсона и Фишера.
8. Распределения с выбросами. Что с ними делать?
9. Робастные методы.
10. БФП. Определение. Теорема Парсеваля. Карта частот.
11. БФП. Свертка. Окна. Расплывание спектра.
12. Спектры и кросс-спектры мощности. Достоверность вычислений.
13. Децимация. Наложение частот, фильтры.
14. АРСС модели. Определение. Вычисление коэффициентов.
15. АРСС фильтры. Устойчивость
16. Сравнение моделей.
17. Статистические оценки через ФМП. Среднее, дисперсия, МНК для нормальной модели.

18. Гауссово приближение к произвольной ФМП.

5. Рекомендованная литература и ссылки по теме

5.1. Основной список

1. П.П. Бочаров. Теория вероятностей и математическая статистика: учебное пособие для вузов. М.: Физматлит, 2005 .— 295 с.
2. П.П. Бочаров. Математическая статистика: учебное пособие для вузов. М.: Изд-во РУДН, 1994.— 165 с.
3. А.В. Печинкин и др. Теория Вероятностей. МГТУ им. Баумана, 2004
4. В.Б. Горяинов и др. Математическая статистика. МГТУ им. Баумана, 2001
5. Дж. Бендат, А. Пирсол. Прикладной анализ случайных данных. М., Мир, 1989

5.2. Дополнительный список

6. А.Д. Наследов, Математические методы психологического исследования, 2007
7. Г. Дженкинс, Д.Ваттс, Спектральный анализ и его приложения, М., Мир, 1971
8. Р.Отнес, Л.Эноксон, Прикладной анализ временных рядов, М., Мир, 1982
9. С.Л. Марпл-мл. Цифровой спектральный анализ и его приложения,1990
10. Э. Айфичер, Б. Джервис, Цифровая обработка сигналов, М. 2004