

Белорусский государственный университет

УТВЕРЖДАЮ

Проректор по учебной работе и
образовательным инновациям

О.И. Чуприс

« 16 »

07

2018 г.

Регистрационный № УД- 6189 /уч.

ПОСЛЕДОВАТЕЛЬНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ И МОНИТОРИНГ ПОТОКОВ ДАННЫХ

**Учебная программа учреждения высшего образования
по учебной дисциплине для специальности второй ступени
высшего образования (магистратуры):**

1-31 81 12 Прикладной компьютерный анализ данных

2018 г.

Учебная программа составлена на основе образовательного стандарта высшего образования ОСВО 1-31 81 12-2015 и учебного плана УВО № G31-251/уч. от 26.05.2017 г.

СОСТАВИТЕЛЬ:

А.Ю. Харин, заведующий кафедрой теории вероятностей и математической статистики Белорусского государственного университета, кандидат физико-математических наук, доцент

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой теории вероятностей и математической статистики Белорусского государственного университета
(протокол № 11 от 27.03.2018 г.);

Научно-методическим Советом Белорусского государственного университета
(протокол № 7 от 13.07.2018 г.)



ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Цель учебной дисциплины «Последовательный статистический анализ и мониторинг потоков данных» – ознакомление студентов II ступени высшего образования (магистрантов) с основными вероятностными моделями и методами, используемыми в последовательном статистическом анализе, и выработка навыков разработки статистических процедур наискорейшего в среднем (по числу наблюдений) принятия решений о параметрах наблюдаемых стохастических потоков данных. При изложении учебной дисциплины важно уделить внимание компьютерной реализации основных методов изучаемой теории.

В рамках поставленной цели **задачи учебной дисциплины** состоят в следующем:

1. Изучение основных вероятностных моделей наблюдений, позволяющих математически описать исследуемые потоки данных.
2. Ознакомление с основными теоретическими результатами в последовательном статистическом анализе.
3. Развитие практических навыков применения изученных моделей методов для мониторинга потоков данных.

Учебная дисциплина «Последовательный статистический анализ и мониторинг потоков данных» относится к компоненту учреждения высшего образования цикла дисциплин специальной подготовки (дисциплина по выбору).

Знания, полученные в результате изучения учебной дисциплины «Последовательный статистический анализ и мониторинг потоков данных», будут использованы при изучении учебной дисциплины II ступени высшего образования «Методы визуализации в анализе больших данных», а также способствовать успешному прохождению производственной практики и подготовки магистерской диссертации.

В результате освоения учебной дисциплины студент магистратуры должен:

знать:

- наиболее часто используемые современные вероятностные модели наблюдаемых стохастических данных;
- основные методы построения последовательных статистических тестов;
- важнейшие свойства характеристик эффективности последовательных статистических процедур и их зависимость от параметров теста;

уметь:

- самостоятельно выбирать вероятностную модель наблюдений в реальной прикладной задаче анализа потока данных;
- формулировать гипотетические предположения и строить последовательный статистический тест проверки сформулированных гипотез;
- вычислять приближенные значения характеристик эффективности используемого последовательного статистического теста;

владеть:

- основными моделями потоков стохастических данных;
- основными методами построения последовательных статистических процедур проверки гипотез о параметрах наблюдаемого потока данных;
- современными информационными технологиями для решения реальных задач мониторинга наблюдаемого потока данных.

Освоение учебной дисциплины «Последовательный статистический анализ и мониторинг потоков данных» должно обеспечить формирование следующих академических, социально-личностных и профессиональных компетенций:

академические компетенции:

АК-1 Способность к самостоятельной профессиональной деятельности (анализ, сопоставление, систематизация, абстрагирование, моделирование, проверка достоверности данных, принятие решений и др.), готовность генерировать и использовать новые идеи.

социально-личностные компетенции:

СЛК-1 Учитывать социальные и нравственно-этические нормы в социально-профессиональной деятельности.

СЛК-2 Быть способным к сотрудничеству и работе в команде.

СЛК-3 Владеть коммуникативными способностями для работы в междисциплинарной и международной среде.

профессиональные компетенции:

ПК-1 Квалифицированно использовать современные достижения по разработке и анализу математических моделей, методов компьютерного анализа данных и современные информационные технологии.

ПК-3 Самостоятельно разрабатывать эффективные численные методы и алгоритмы, а также интегрировать их в компьютерные системы анализа данных.

Структура содержания учебной дисциплины включает такие дидактические единицы, как темы и разделы, в соответствии с которыми разрабатываются и реализуются соответствующие лекционные и лабораторные занятия. Примерная тематика лабораторных работ приведена в информационно-методической части.

Дисциплина изучается в третьем семестре (II ступень). Всего на изучение учебной дисциплины «Последовательный статистический анализ и мониторинг потоков данных» отведено 112 часов, в том числе 40 аудиторных часов, из них: лекции – 20 часов, лабораторные занятия – 20 часов.

Трудоемкость учебной дисциплины составляет 3 зачетные единицы.

Форма текущей аттестации – зачет.

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

Раздел 1. Особенности последовательного статистического анализа

Тема 1.1. Введение, основные понятия и определения

Введение. Основное отличие последовательного анализа от статистических методов, основанных на фиксированном объеме выборки. Примеры. Основные определения.

Тема 1.2. Характеристики эффективности, их вычисление

Характеристики эффективности последовательного статистического теста на примере последовательного критерия отношения вероятностей. Проблема вычисления и методы ее решения.

Раздел 2. Последовательный статистический анализ для различных моделей данных

Тема 2.1. Последовательность независимых одинаково распределенных наблюдений

Случай простых гипотез. Построение теста для проверки сложных гипотез.

Тема 2.2. Неоднородные наблюдения

Случай неоднородных независимых наблюдений. Построение последовательного теста для наблюдений, образующих временной ряд с трендом. Усечение последовательной процедуры по числу наблюдений

Тема 2.3. Наблюдения, образующие цепь Маркова

Построение последовательного теста проверки гипотез о параметрах наблюдаемой цепи Маркова. Анализ характеристик эффективности.

Раздел 3. Мониторинг потоков данных для выявления скачкообразных изменений («разладок»)

Тема 3.1. Постановка задачи о «разладке». Применение последовательного анализа для решения поставленной задачи

Основные типовые случаи (модели) «разладок», встречающиеся на практике при анализе потоков стохастических данных. Методика применения последовательного подхода для выявления факта и момента «разладки».

Тема 3.2. Метод кумулятивных сумм CUSUM

Алгоритм метода. Математические свойства. Простейшие применения.

Тема 3.3. Применения метода CUSUM

Применения метода кумулятивных сумм для выявления типовых скачкообразных изменений значений параметров наблюдаемого потока данных.

Тема 3.4. Современные алгоритмы выявления «разладок», основанные на последовательных статистических тестах

Современные методы выявления «разладок» с последовательностями случайных величин, описываемых специальными моделями.

Тема 3.5. Современные проблемы последовательных методов статистического анализа и мониторинга потоков данных

Современные библиотеки компьютерных процедур для реализации последовательных методов статистического анализа и мониторинга потоков данных. Наиболее актуальные теоретические и прикладные задачи. Заключение.

УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

Номер раздела, темы	Название раздела, темы	Количество аудиторных часов					Количество часов УСР	Форма контроля знаний
		Лекции	Практические занятия	Семинарские занятия	Лабораторные занятия	Иное		
1	2	3	4	5	6	7	8	9
1	Особенности последовательного статистического анализа	4			4			
1.1	Введение, основные понятия и определения	2			2			Устный опрос
1.2	Характеристики эффективности, их вычисление	2			2			Коллоквиум
2	Последовательный статистический анализ для различных моделей данных	6			6			
2.1	Последовательность независимых одинаково распределенных наблюдений	2			2			Устный опрос
2.2	Неоднородные наблюдения	2			2			Устный опрос
2.3	Наблюдения, образующие цепь Маркова	2			2			Отчет по лаб. работе №1
3	Мониторинг потоков данных для выявления скачкообразных изменений («разладок»)	10			10			
3.1	Постановка задачи о «разладке». Применение последовательного анализа для решения поставленной задачи	2			2			Устный опрос
3.2	Метод кумулятивных сумм CUSUM	2			2			Устный опрос
3.3	Применения метода CUSUM	2			2			Устный опрос
3.4	Современные алгоритмы выявления «разладок», основанные на	2			2			Командная игра для твор-

	последовательных статистических тестах							ческого решения заданий открытого типа
3.5	Современные проблемы последовательных методов статистического анализа и мониторинга потоков данных	2			2			Отчет по лаб. работе №2
	Всего: 40 часов	20			20			

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

Перечень основной литературы

1. Ширяев, А.Н. Статистический последовательный анализ. Оптимальные правила остановки / А.Н. Ширяев. – 2-е изд. – М.: Editorial URSS, 2018. – 272 с.

2. Харин, А.Ю. Робастность байесовских и последовательных статистических решающих правил / А.Ю. Харин. – Минск : БГУ, 2014. – 207 с.

Перечень дополнительной литературы

1. Харин Ю.С. Теория вероятностей, математическая и прикладная статистика / Ю.С. Харин, Н.М. Зуев, Е.Е. Жук. – Минск : БГУ. – 2011. – 463 с.

2. Вальд А. Последовательный анализ: Пер. с англ. / Севастьянов Б.А., ред.- М.: Физматлит, 1960. - 328 с.

3. Ghosh, B.K. Handbook of sequential analysis / B.K. Ghosh, P.K. Sen. – New York : Marcel Dekker, 1991. - 1012 p.

Перечень рекомендуемых средств диагностики

Для текущего контроля качества усвоения знаний студентами используется следующий диагностический инструментарий:

- устные опросы;
- командная игра для творческого решения заданий открытого типа;
- проведение коллоквиума;
- защита отчетов по лабораторным работам.

Методика формирования итоговой оценки

Итоговая оценка формируется на основе:

1. Правил проведения аттестации студентов (Постановление Министерства образования Республики Беларусь № 53 от 29 мая 2012 г.);

2. Положения о рейтинговой системе оценки знаний по дисциплине в БГУ (Приказ ректора БГУ от 18.08.2015 № 382-ОД);

3. Критериев оценки знаний студентов (письмо Министерства образования от 22.12.2003 г.).

В течение семестра проводятся коллоквиум и две лабораторные работы. На выполнение коллоквиума отводится до 30 минут. Задания оцениваются в соответствии с их сложностью, максимальная сумма баллов за все задания в коллоквиуме равна 10.

Итоговая оценка по дисциплине рассчитывается как средневзвешенная оценка коллоквиума и лабораторных работ. Весовой коэффициент для оценки коллоквиума – 0,2; для оценки каждой лабораторной работы – 0,4.

Примерная тематика лабораторных работ

Лабораторная работа № 1. Компьютерная реализация и анализ эффективности последовательного статистического теста для конкретной модели наблюдений.

Лабораторная работа № 2. Компьютерная реализация и исследование алгоритма кумулятивных сумм CUSUM для конкретной задачи анализа данных.

ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ УВО

Название учебной дисциплины, с которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы учреждения высшего образования по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола) ¹
Статистический анализ неполных данных	Теории вероятностей и математической статистики	Нет	Оставить содержание учебной дисциплины без изменения, протокол № 11 от 27.03.2018 г.

ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ ПО ИЗУЧАЕМОЙ УЧЕБНОЙ ДИСЦИПЛИНЕ

на ____ / ____ учебный год

№ п/п	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры
 _____ (название кафедры) (протокол № ____ от _____ 201_ г.)

Заведующий кафедрой

_____ (ученая степень, ученое звание)

_____ (подпись)

_____ (И.О. Фамилия)

УТВЕРЖДАЮ
Декан факультета

_____ (ученая степень, ученое звание)

_____ (подпись)

_____ (И.О. Фамилия)