

**Федеральное государственное автономное образовательное учреждение  
высшего образования  
"Национальный исследовательский университет  
"Высшая школа экономики"**

Факультет Компьютерных наук  
Департамент больших данных и информационного поиска  
Базовая кафедра Яндекс

**Рабочая программа дисциплины «Вероятностно-статистический методы в  
теории принятия решений»**

для образовательной программы «Науки о данных»  
направления подготовки 01.04.02 "Прикладная математика и информатика"  
уровень магистра

Разработчик(и) программы  
Пузыревский И.В. (ipuzurevskiy@hse.ru)

Одобрена на заседании базовой кафедры Яндекс  
«\_\_»\_\_\_\_\_ 2017 г.

Заведующий Кафедрой  
М.А. Бабенко \_\_\_\_\_

Утверждена Академическим советом образовательной программы  
«\_\_»\_\_\_\_\_ 2017 г., № протокола \_\_\_\_\_

Академический руководитель образовательной программы  
С.О. Кузнецов \_

Москва, 2017

*Настоящая программа не может быть использована другими подразделениями  
университета и другими вузами без разрешения подразделения-разработчика программы*



### 1. Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает требования к образовательным результатам и результатам обучения студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих дисциплину Научный семинар «Вероятностно-статистические методы в теории принятия решений», учебных ассистентов и студентов направления подготовки 01.04.02 «Прикладная математика и информатика», обучающихся по образовательной программе «Науки о данных».

Программа учебной дисциплины разработана в соответствии с:

- Образовательным стандартом федерального государственного автономного образовательного учреждения высшего профессионального образования «Национального исследовательского университета «Высшая школа экономики»;
- Образовательной программой подготовки магистра по направлению 01.04.02 «Прикладная математика и информатика», специализации «Анализ Интернет-данных».
- Объединенным учебным планом университета по образовательной программе «Науки о данных», утвержденным в 2018 г.

### 2. Место дисциплины в структуре образовательной программы

Предлагаемый курс позволит студентам с одной стороны ознакомиться с основными теоретическими свойствами алгоритмов скорейшего обнаружения разладки, а с другой стороны познакомит их с особенностями практической реализации и примерами применения современных алгоритмов скорейшего обнаружения разладки и детектирования аномалий.

### 3. Цели освоения дисциплины

Целью данного курса является ознакомление слушателей с общей теорией оптимальной остановки стохастических процессов, научить их строить оптимальные методы обработки поступающей и поступившей информации, в том числе и методы скорейшего обнаружения разладки. Кроме того, слушатели должны получить понятие об использовании методов скорейшего обнаружения разладки и детектирования аномалий для решения практически задач в различных областях науки и техники.

### 4. Тематический план дисциплины «Вероятностно-статистические методы в теории принятия решений»

№	Название темы	Всего часов по дисциплине	Аудиторные часы		Самостоятельная работа
			Лекции	Сем. и практика	
1	Основы теории скорейшего обнаружения разладки	24	6	8	10



2	Практическая реализация стандартных статистик для скорейшего обнаружения разладки	34	6	8	20
3	Некоторые избранные методы обработки и моделирования временных рядов	24	6	8	10
4	Модификации стандартных методов скорейшего обнаружения для случая неизвестных параметров сигнала после разладки	30	4	6	20
5	Обнаружение аномалий	32	4	8	20
	Итого	144	26	38	80

### 5. Формы контроля и структура итоговой оценки

Текущий контроль - домашняя работа в первом модуле, контрольная работа в первом модуле.

Итоговый контроль – письменный экзамен (120 мин.)

Итоговая оценка вычисляется следующим образом:

$0,1 \cdot \text{оценка за домашнюю} + 0,2 \cdot \text{оценка за контрольную} + 0,7 \cdot \text{оценка за экзамен.}$

**Таблица соответствия оценок по десятибалльной и системе зачет/незачет**

Оценка по 10-балльной шкале	Оценка по 5-балльной шкале
1	Незачет
2	
3	
4	Зачет
5	
6	
7	
8	
9	
10	

**Таблица соответствия оценок по десятибалльной и пятибалльной системе**

По десятибалльной шкале	По пятибалльной системе
1 – неудовлетворительно	неудовлетворительно – 2
2 – очень плохо	
3 – плохо	
4 – удовлетворительно	удовлетворительно – 3
5 – весьма удовлетворительно	



6 – хорошо 7 – очень хорошо	хорошо – 4
8 – почти отлично 9 – отлично 10 – блестяще	отлично – 5

## 6. Программа дисциплины «Вероятностно-статистические методы в теории принятия решений»

### Тема 1. Основы теории скорейшего обнаружения разладки

- Обзор основных задач в теории принятия решений
- Ключевые статистики и тесты в теории принятия решений в задачах различения двух гипотез по фиксированному числу наблюдений. Дискретное время
- Ключевые статистики и последовательные тесты в задачах различения двух гипотез.
- Дискретное время
- Некоторые широко используемые статистики в проблеме скорейшего обнаружения
- Об определении броуновского движения, производной Радона-Никодима и условного математического ожидания
- О различении двух гипотез для броуновского движения. Сравнение методов Неймана-
- Пирсона и Вальда
- Об основных формулировках задач скорейшего обнаружения для броуновского движения (Варианты А, В, С, D и E)
- "Вальдовская" система наблюдения в варианте E
- "Нейман-Пирсоновская" система наблюдения в варианте B
- Постановка задачи о "разладки" в варианте A и ее редукция к задаче об оптимальной остановке для марковского процесса
- Решение задачи об оптимальной остановке для марковского процесса в варианте A

### Тема 2. Практическая реализация стандартных статистик для скорейшего обнаружения разладки

- Контрольные карты Шухарта
- CUSUM, "двухсторонний" CUSUM
- Ширяев-Робертс
- Модификация стандартных статистик для выявления разладки в дисперсии и ковариационной структуре (случай многомерного сигнала)

### Тема 3. Некоторые избранные методы обработки и моделирования временных

## **рядов**

- Моделирование временных рядов на основе моделей типа ARMA
- Обработка сигналов: очистка от шума, сегментация сигналов на основе скрытых марковских моделей, сглаживания с адаптивными весами
- Экспоненциальное сглаживание: оценки среднего значения и параметра масштаба, экспоненциальное сглаживание в случае неравномерного расположения наблюдений, метод Хольта-Винтерса, робастный вариант экспоненциального сглаживания

## **Тема 4. Модификации стандартных методов скорейшего обнаружения для случая неизвестных параметров сигнала после разладки**

- Практическая реализация стандартных методов скорейшего обнаружения. Статистическое моделирование распределения ложных тревог
- Байесовские предположения о параметрах распределения после разладки
- Использование ОМП для задания значений параметров распределения после разладки
- Мгновенная оценка среднего значения после разладки на основе текущего наблюдения
- Оценка среднего значения и параметра масштаба на основе экспоненциального сглаживания
- Численное оценивание среднего времени до ложной тревоги и среднего времени до обнаружения разладки для стандартных алгоритмов скорейшего обнаружения
- Настройка параметров алгоритмов скорейшего обнаружения

## **Тема 5. Обнаружение аномалий**

- Основные методы обнаружения аномалий и их теоретические свойства
- Примеры использования методов наискорейшего обнаружения и детектирования аномалий

## **7. Учебно-методическое и информационное обеспечение дисциплины Список литературы**

### **Основная литература**

1. Вальд А. Последовательный анализ: Пер. с англ. / Севастьянов Б.А., ред.- М.: Физматлит, 1960. - 328 с.
2. Леман Э. Проверка статистических гипотез. Изд-во "Наука", М., 1964.
3. Ширяев А.Н. Статистический последовательный анализ: Оптимальные правила остановки. - 2-е изд., перераб. - М.: Наука. Физматлит, 1976. - 272 с. - (Теория вероятностей и математическая статистика).

4. Ширяев А.Н. Вероятность-1,2. 4-ое издание. М.:МЦМНО, 2007.
5. Н. Клигине, Л. Телькснис. Методы обнаружения моментов изменения свойств случайных процессов (обзор) // Автоматика и телемеханика. 1983. № 10. С. 5-56.
6. И.В. Никифоров Последовательное обнаружение изменения свойств временных рядов. М.: Наука, 1983.
7. М. Бассевиль, А. Банвенист. Обнаружение изменения свойств сигналов и динамических систем. М.: Наука, 1989.
8. А.Г. Тартаковский. Последовательные методы в теории информационных систем. М.: Радио и связь, 1991.
9. Peskir Goran, Shiryaev Albert. Optimal stopping and free-boundary problems. Basel etc.: Birkhauser-Verl., 2006. - 500 с. - Bibliogr.: p. 477-492.
10. M. Basseville, I. V. Nikiforov. Detection of abrupt changes: theory and application. Prentice-Hall, 1993.
11. D. Siegmund Sequential analysis : Tests and confidence intervals.- New York etc.: Springer-Verl., 1985. - 272 pp. - (Springer series in statistics). - Bibliogr.: pp.263-270. - ISBN 0-387-96134-8.
12. A.S. Willsky. A survey of design methods for failure detection in dynamic systems // Automatica. 1976. V. 12. P. 601-611. 2
13. B.K. Ghosh, P.K. Sen. Handbook of sequential analysis. Marcel Dekker, Inc., 1991.
14. E. Carlstein, H.-G. Muller, D. Siegmund (eds.). Change-point problems. Papers presented at the AMS-IMS-SIAM Summer Research Conference on Change-Point Problems, Mt. Holyoke College, July 11-17, 1992.