

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Родионова Светлана Евгеньевна
Должность: Начальник учебно-методического управления
Дата подписания: 14.12.2021
Уникальный программный ключ:
3d7c75ac99fd0ac390d8867fe19b94e675a67209f5692fc73e4e4767f4223223

I. ОБРАЗОВАТЕЛЬНАЯ ПРОГРАММА

Федеральное государственное бюджетное образовательное учреждение высшего образования

«БАШКИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

УТВЕРЖДАЮ

Ректор

« 5 »

Н.Д. Морозкин



20 12 21

Дополнительная профессиональная программа повышения квалификации

«Предиктивная аналитика на основе анализа временных рядов»

72 час.

СОГЛАСОВАНО

Директор ИНО  Т.Б.Великханина

« 5 » 10 2020г

УФА 2020

ОБЩАЯ ХАРАКТЕРИСТИКА ПРОГРАММЫ

1. Цель программы дать систематизированное представление о современных подходах к анализу данных, в том числе больших данных, представленных временными рядами, научить применению основных принципов и этапов решения задач построения адекватных моделей прогнозирования с использованием современной информационной среды (R Studio), развить умения проверки прогнозных свойств модели прогнозирования, в конечном итоге сформировать на продвинутом уровне компетенцию цифровой экономики: способность строить обоснованные прогнозы на основе анализа временных рядов.

2. Планируемые результаты обучения:

а. Знание (осведомленность в областях)

- 2.1.1 типы процессов, представленные временными рядами (TSP, DSP);
- 2.1.2 детерминированные составляющие временного ряда и методы очищения от них случайных процессов;
- 2.1.3 тесты на единичные корни и их модификации;
- 2.1.4 детерминированные компоненты временного ряда
- 2.1.5 модели ARIMA/SARIMA/ARIMAX;
- 2.1.6 тесты на структурные изменения временных рядов;
- 2.1.7 модели адаптивного прогнозирования (экспоненциального сглаживания);
- 2.1.8 схему обобщённого прогнозирования на основе моделирования случайных процессов и методы оценки прогностической ценности модели.

2.2 Умение (способность к деятельности):

- 2.2.1 определять структуру временного ряда на основе анализа коррелограмм ACF и PACF;
- 2.2.2 определять типы процессов, представленные временными рядами на основе различных процедур (например, с использованием пятиэтапной процедуры применения теста ADF);
- 2.2.3 идентифицировать порядки, проводить селекцию, оценку и диагностику моделей ARIMA, SARIMA
- 2.2.4 Строить модели анализа временных рядов с введением экзогенной переменной (ARIMAX)
- 2.2.5 Использовать модели адаптивного сглаживания для краткосрочного прогнозирования
- 2.2.6 строить научно-обоснованные прогнозы, проводить корректную интерпретацию результатов прогнозирования

2.3 Навык (использование конкретных инструментов):

- 2.3.1 навыками проведения анализа внутренней структуры различных процессов, представленных временными рядами.
- 2.3.2 построения адекватных моделей ARIMA, SARIMA, ARIMAX в среде R Studio
- 2.3.3 навыками оценки прогнозных свойств полученных моделей
- 2.3.4 навыками построения качественных прогнозов с использованием методов математического моделирования случайных процессов для корректных принятий управленческих решений

- 3 Требования к слушателям** (возможно заполнение не всех полей)
- 3.1 Образование: высшее, среднее профессиональное
- 3.2 Квалификация: инженер, математик, экономист
- 3.3 Наличие опыта профессиональной деятельности: работа в Excel.
- 3.4 Предварительное освоение иных дисциплин/курсов /модулей: высшая математика, теория вероятностей и математическая статистика, общая теория статистики

4. Учебный план программы «Предиктивная аналитика на основе анализа временных рядов»

№ п/п	Модуль	Всего, час	Виды учебных занятий		
			лекции	практические занятия	самостоятельная работа
0	Входное тестирование	1		1	
1	Модуль 1. Введение в курс	7	2	2	3
2	Модуль 2. Типы случайных процессов, представленные временными рядами	16	4	6	6
3	Модуль 3. Модели нестационарных случайных процессов ARIMA/SARIMA/ARIMAX	18	6	6	6
4	Модуль 4. Модели адаптивного краткосрочного прогнозирования	12	3	4	4
5	Модуль 5 Обобщённое прогнозирование	10	2	3	5
Итоговая аттестация			зачет		
Итоговое задание		8	Зачет - Кейс-проект		

5. Календарный план-график реализации образовательной программы

(дата начала обучения – дата завершения обучения) в текущем календарном году, указания на периодичность набора групп (не менее 1 группы в месяц)

№ п/п	Наименование учебных модулей	Трудоёмкость (час)	Сроки обучения
0	Входное тестирование	1	1.11.2020
1	Модуль 1. Введение в курс	7	2.11.2020-3.11.2020-
2	Модуль 2. Типы случайных процессов, представленные временными рядами	16	4.11.2020-7.11.2020-
3	Модуль 3. Модели нестационарных случайных процессов ARIMA/SARIMA/ARIMAX	18	8.11.2020-12.11.2020
4	Модуль 4. Модели адаптивного краткосрочного прогнозирования	12	13.11.2020-15.11.2020

5	Модуль 5 Обобщённое прогнозирование	10	16.11.2020-18.11.2020
	Итоговое задание	8	19.11.2020-21.11.2020
Всего:		72	1.11.2020-21.11.2020

6. Учебно-тематический план программы «Анализ данных на основе машинного обучения»

№ п/п	Модуль / Тема	Всего, час	Виды учебных занятий			Формы контроля
			лекции	практические занятия	самостоятельная работа	
0	Входное тестирование	1			1	Тест
1	Модуль 1. Введение в курс	7	2	2	3	Кейс 1
2	Модуль 2. Типы случайных процессов, представленные временными рядами	16	4	6	6	Тест
2.1	Определение типа случайного процесса	8	2	3	3	Кейс 2
2.2	Уточнение процесса с учётом структурных изменений в динамике ряда	8	2	3	3	Кейс 3
3	Модуль 3. Модели нестационарных случайных процессов ARIMA/SARIMA/ARIMAX	18	6	6	6	Тест
3.1.	Оценка моделей ARIMA	6	2	2	2	Кейс 4
3.2.	Оценка сезонности, SARIMA	6	2	2	2	Кейс 5
3.3	Введение экзогенных переменных в модель ARIMAX	6	2	2	2	Кейс 6
4	Модуль 4. Модели адаптивного краткосрочного прогнозирования	12	3	4	4	Кейс 7 Тест
5	Модуль 5 Обобщённое прогнозирование	10	2	3	5	Кейс 8
8	Итоговая аттестация	8			8	Кейс-итоговый проект

7. Учебная (рабочая) программа повышения квалификации « Анализ данных на основе машинного обучения»

7.1 Модуль 1 «Введение в курс» (7 ак.часов)

Темы

Общие требования, предъявляемые к временным рядам. Задачи, требующие решения на основе анализа временных рядов. Различия в краткосрочные и долгосрочные прогнозирование. Принципы для составления информационных требований к базе прогнозирования. Постановка задач обучения для временных рядов.

Задания в виде кейса 1 Установка R Studio, определение настроек.

7.2 Модуль 2 «Типы случайных процессов, представленные временными рядами» (16 ак. часов)

Темы

Понятие детерминированного тренда и процесса случайного блуждания, в том числе с дрейфом. Стационарность случайных стохастических процессов в широком и узком смысле. Разделение DS и TS процессов. Интегрируемые стохастические процессы, порядок интегрируемости. Тест Дики-Фуллера. Определение порядка интегрируемости. Расширенный тест Дики-Фуллера. Подход Доладо-Дженкинсона-Сосвила-Риверо. Процесс белого шума. Методы определения детерминированного тренда в структуре временного ряда. Определение структурных изменений в динамике ряда, разделение TS и Ds с учётом структурных изменений. Тесты на структурные изменения временных рядов: Тест Перрона, Тест Чоу с заранее известной точкой излома, Тест Квандта-Эндрюса.

Задания в виде кейса 2 и кейса 3 (определение типа случайного процесса, представленного временным рядом, с учётом структурных изменений в динамике), тест.

7.3 Модуль 3 Модели нестационарных случайных процессов ARIMA/SARIMA/ARIMAX (18 ак. часов)

Темы

Выделение сезонной компоненты (аддитивная и мультипликативная модели). Оценка сезонной компоненты с помощью тригонометрических функций. Оценка сезонной компоненты методом сезонных индексов. Оценка сезонной компоненты методом фиктивных переменных. Определение авторегрессионных (AR) процессов. Модели скользящих средних (MA). Авторегрессионные (ARMA) модели скользящей средней. Автокорреляционная функция (АКФ) и ее свойства. Частная автокорреляционная функция (ЧАКФ) и ее свойства. Критерий для ARMA процессов Люнга –Бокса. Идентификация модели ARMA по коррелограммам АКФ и ЧАКФ. Проверка адекватности построенной ARMA -модели. ARIMA-модели. Подход Бокса-Дженкинса. Идентификация моделей. Сезонные ARIMA-модели (SARIMA). Селекция моделей на основе информационных критериев. Введение экзогенной переменной в модель ARIMA – ARIMAX.

Задания в виде кейсов 4, 5, 6 (построение моделей ARIMA/SARIMA/ARIMAX). Тест.

7.4 Модуль 4 Модели адаптивного краткосрочного прогнозирования (12 ак. часов)

Темы

Общие сведения о моделях экспоненциального сглаживания. Простая линейная модель Брауна. Методы подбора адаптирующих параметров. Квадратичная модель Брауна. Модель Хольта. Адаптивные модели с учетом сезонности: Мультипликативная и аддитивная модели Хольта-Уинтерса. Модель Тейла-Вейджи. Проверка адекватности построенных прогнозных адаптивных моделей.

Задания в виде кейса 7 (построение адекватных моделей адаптивного прогнозирования). Тест.

7.5 Модуль 5 Обобщённое прогнозирование (10 ак. часов)

Темы

Схема составления прогнозных моделей. Информационная база прогнозирования. Прогнозирование на основе динамических моделей. Доверительные интервалы в прогнозах. Проверка прогностической

ценности прогнозов. Метрики качества построения прогнозов. Особенности валидации в моделировании на временных рядах.

Задания в виде кейса 8 (составление обобщённого прогноза).

Описание практико-ориентированных заданий и кейсов

	Номер темы/модуля	Наименование практического занятия	Описание
1.1	1	Кейс-задание 1. Установка R Studio	Цель практического задания: формирование навыков работы с R Studio, установка соответствующих настроек
1.2.	2.	Кейс-задание 2. Тестирование случайных процессов, представленные временными рядами, для определения его типа	Цели практического задания: научиться тестировать временной ряд на тип процессов
1.3	2	Кейс-задание 3. Определение типов случайных процессов, представленные временными рядами, с учётом структурных изменений	Цели практического задания: научиться тестировать временной ряд на тип процессов с учётом структурных изменений в его динамике
1.4	3	Кейс-задание 4 (модуль 3). Построение моделей нестационарных случайных процессов ARIMA	Цели практического задания: построение модели ARIMA и оценка её адекватности реальному моделируемому процессу.
1.5	3	Кейс-задание 5. (Модуль 3). Построение моделей нестационарных случайных процессов SARIMA	Цели практического задания: построение модели SARIMA и оценка её адекватности реальному моделируемому процессу
1.6.	3	Кейс-задание 6. (модуль 3). Построение моделей нестационарных случайных процессов ARIMAX	Цели практического задания: построение модели ARIMAX и оценка её адекватности реальному моделируемому процессу.
1.7	4	Кейс-задание 7. Адаптивные методы краткосрочного прогнозирования	Цели практического задания: построение адаптивной модели прогнозирования и оценка её адекватности реальному моделируемому процессу
1.8	5	Кейс-задание 8. Обобщённое прогнозирование	Цели практического задания: построение прогнозов согласно полученным моделям и оценка их качества

8. Оценочные материалы по образовательной программе

8.1. Вопросы тестирования по модулям

№ модуля	Вопросы входного тестирования	Вопросы промежуточного тестирования	Вопросы итогового тестирования
0	<p>1. Сумма двух событий – это событие, состоящее</p> <ul style="list-style-type: none">A. в одновременном их появленииB. в появлении по крайней мере одного из нихC. в их последовательном появленииD. в не появлении ни одного из них <p>2. Множество результатов, отобранных из генеральной совокупности, называют</p> <ul style="list-style-type: none">A. ВыборкойB. Вариационным рядомC. Статистикой критерияD. Точечными оценками <p>3. Статистическая гипотеза – это утверждение о свойствах</p> <ul style="list-style-type: none">A. генеральной совокупностиB. выборкиC. конкретного объекта <p>4. Значение признака, находящееся в середине вариационного ряда наблюдений,</p> <ul style="list-style-type: none">a. мода;b. средняя арифметическая;c. медиана;d. частота;e. частость. <p>5. Названия гипотезы, противоположной проверяемой:</p> <ul style="list-style-type: none">A. нулеваяB. простаяC. конкурирующаяD. альтернативнаяE. Сложная <p>6. Интервал возможных значений парного коэффициента корреляции при наличии между величинами X и Y отрицательной, но не функциональной связи:</p> <ul style="list-style-type: none">A. (-1; 0)B. (0; 1)C. (-1; -0,5)D. (-0,5; 0)E. [-1; 0]		

	<p>7. Под трендом понимают:</p> <p>А. Краткосрочное развитие</p> <p>В. Сдвиг по времени</p> <p>С. Долговременную тенденцию развития</p> <p>Д. Периодическую тенденцию развития</p> <p>8. Ретроспективная история ряда содержит ежемесячные наблюдения. Период сезонности составляет:</p> <p>А. 6</p> <p>В. 4</p> <p>С. 12</p> <p>Д. 24</p> <p>9. Для построения прогнозных моделей финансовых показателей требуется предварительно:</p> <p>А. Перевести показатели временных рядов в сопоставимые цены</p> <p>В. Перевести показатели временных рядов к одному базовому периоду</p> <p>С. Перевести показатели временных рядов к предыдущему периоду</p> <p>Д. Никаких действий выполнять не требуется</p>		
2		<ol style="list-style-type: none"> 1. Понятие временного ряда. Общие положения о временных рядах. Определение временного ряда через случайные стохастические процессы. 2. Понятие детерминированного тренда и процесса случайного блуждания, в том числе с дрейфом. Процесс белого шума. Теорема разложения Вольда. 3. Стационарность случайных стохастических процессов в широком и узком смысле. Разделение DS и TS процессов. Подход Доладо-Дженкинсона-Сосвила-Риверо. 4. Интегрируемые стохастические процессы, порядок интегрируемости. Тест Дикки-Фуллера. Определение порядка интегрируемости. Расширенный тест Дики-Фуллера. 	

		<ol style="list-style-type: none"> 5. Тест на сезонную интегрируемость Дики, Хаза, Фуллера. Тест Филиппса-Перрона. 6. Общие положения о компонентном анализе временных рядов. 7. Методы определения детерминированного тренда в структуре временного ряда. 8. Виды структурных изменений временных рядов (скачки, изломы). 9. Тесты на структурные изменения временных рядов: Тест Перрона, Тест Чоу с заранее известной точкой излома, Тест Рамсея, Тест Квандта-Эндрюса. 10. Методы избавления от структурных изломов: подход Гуаратти. 	
3		<ol style="list-style-type: none"> 1. Определение авторегрессионных (AR) процессов. Модели скользящих средних (MA). Авторегрессионные (ARMA) модели скользящей средней. 2. Автокорреляционная функция (АКФ) и ее свойства. Частная автокорреляционная функция (ЧАКФ) и ее свойства. 3. Критерий для ARMA процессов Люнга –Бокса. Идентификация модели ARMA по коррелограммам АКФ и ЧАКФ. 4. Проверка адекватности построенной ARMA -модели. 5. ARIMA-модели. Подход Бокса-Дженкинса. Идентификация моделей. 6. Сезонные ARIMA-модели (SARIMA). 7. Селекция моделей ARIMA на основе информационных критериев. 8. Определение взаимосвязи временных рядов на основе теории коинтеграции, коинтеграционное соотношение. 9. Причинность по Гренджеру, тест Ингла-Гренджера. 10. Модель ARIMAX 	
4		<ol style="list-style-type: none"> 1. Адаптивные методы прогнозирования, общие положения и их классификация. 	

		<p>Определение параметров адаптации.</p> <ol style="list-style-type: none"> 2. Экспоненциальные средние первого, второго и третьего порядков. 3. Адаптивная линейная модель Брауна. 4. Адаптивная квадратичная модель Брауна. 5. Метод подбора параметров адаптации, исходя из минимума средней квадратичной ошибки 6. Метод «Сетка поиска» подбора параметров адаптации 7. Адаптивная модель Хольта. 8. Адаптивная мультипликативная модель Хольта-Уинтерса. 9. Адаптивная модель Тейла-Вейджи 10. Метод Брауна для подбора параметров адаптации 	
5		<ol style="list-style-type: none"> 1. Доверительные интервалы в прогнозах. 2. Проверка прогностической ценности прогнозов: на основе стандартного отклонения. 3. Долевые показатели оценки прогностической ценности прогнозов 4. Коэффициент Тейла для оценки прогностической ценности прогнозов 5. Отличие в построении статического и динамического прогноза 	

8.2. Минимальным проходным баллом теста считается 60% верных ответов по результатам суммарно 2 попыток

Кейс в 10 баллов: Максимум 10 баллов - выставляется при выполнении всех требований к отчету, подробном описании всех этапов и представлении выводов, увязывающих выполненный кейс с встречающимися в практике задачами.

8-10 баллов выставляется при выполнении всех требований к отчету, подробном описании всех этапов или представлении выводов, увязывающих выполненный кейс с встречающимися в практике задачами

6-7 баллов выставляется при выполнении всех требований к отчету, подробном описании отдельных этапов или кратких выводов.

Минимально допустимый балл 5 баллов -выставляется при выполнении минимального требования к отчету кейса

Кейс в 20 баллов: Максимум 20 баллов - выставляется при выполнении всех требований к отчету, подробном описании всех этапов и представлении выводов, увязывающих выполненный кейс с встречающимися в практике задачами.

16-20 баллов выставляется при выполнении всех требований к отчету, подробном описании всех этапов или представлении выводов, увязывающих выполненный кейс с встречающимися в практике задачами

11-15 баллов выставляется при выполнении всех требований к отчету, подробном описании отдельных этапов или кратких выводов.

Минимально допустимый балл 10 баллов -выставляется при выполнении минимального требования к отчёту кейса.

8.3. итоговое задание по всей образовательной программе

Цели задания: оценка сформированности компетенции по предиктивной аналитике на основе анализа временных рядов (способность управлять информацией и знаниями) на продвинутом уровне.

Итоговое комплексное Задание

В итоговом задании необходимо использовать данные, предоставленные в файле Complex_dataset.rds.

1.1 Создайте новый скрипт RMarkdown.

1. Загрузите датасет Complex_dataset.rds

2. Разделите выборки на train/test с соотношением 0.8 с помощью технологии

3. Зависимая (целевая) переменная, подлежащая прогнозированию: Target

1.2 Моделирование

1. Вам необходимо построить три модели:

1. Стандартная модель ARIMA | SARIMA (в зависимости от определения наличия эффекта сезонности в данных), предварительно идентифицировав порядки моделей на основе анализа коррелограмм

2. Модель с введением экзогенной переменной SARIMAX, предварительно идентифицировав порядки моделей на основе анализа коррелограмм и кросскоррелограмм

3. Модель адаптивного сглаживания, предварительно определив параметры адаптации по технологии сетки поиска

2. Проверить адекватность моделей и устойчивость результатов моделирования

3. Построить прогнозы на тестовый период по каждой из модели и оценить метрики:

Квадратный корень средней ошибки предсказания:

- Среднюю ошибку прогноза по модулю

- Среднюю процентную ошибку по модулю (среднюю ошибку аппроксимации)

- Коэффициент неравенства Тейла;

- Долю систематической ошибки;

- Долю вариации;

- Долю ковариации.

2 Оформление отчета

Отчет необходимо оформить в RMarkdown и скомпилировать в HTML или Word. Отключите warning и message в куске кода (чанке) с подключением пакетов. Скрывать код (echo) из отчета не нужно.

8.4. Задания-кейсы

Кейс-задание 1 (по модулю 1). Установка R Studio

1. Скачать необходимые программы для своей операционной системы 1.1. R • Для win – <https://cran.r-project.org/bin/windows/base/>

• Для macos – <https://cran.r-project.org/bin/macosx/>

1.2. RStudio • <https://rstudio.com/products/rstudio/download/#download>

- 1.3. Rtools (только для win) • <https://cran.r-project.org/bin/windows/Rtools/>
2. Установить скаченные программы. 2.1. Порядок установки R -> RStudio -> Rtools
- 2.2. Для пользователей Windows: • Важно, чтобы путь установки НЕ содержал русских букв
 - Нежелательно устанавливать программы в стандартную папку «Program Files», поскольку в дальнейшем возникнут проблемы при установке дополнительных библиотек
 - Лучше всего установить R в «C:/R», RStudio в «C:/RStudio», а Rtools в «C:/Rtools»
3. RStudio – это оболочка для языка R. В данном курсе вы будете использовать именно эту программу. После установки запустите RStudio и убедитесь, что она правильно установлена. Также необходимо изменить некоторые стандартные настройки. Для этого откройте раздел Tools – Global options
 - 3.1. В разделе General установите настройки на UTF8.:

Кейс-задание 2 (по модулю 2). Тестирование случайных процессов, представленные временными рядами, для определения его типа

Согласно процедуре Доладо-Дженкинсона-Сосвила-Риверо определить к какому типу процесса относится временной ряд (при 1, 5 и 10 % уровне значимости), предложенный преподавателем и выбранный слушателем, имеется ли в процессе детерминированный и/или стохастический тренд, определить TS, DS (I(1), I(2) или I(0)), TS+DS процесс определяет структура ряда. Для выполнения задания использовать пятиэтапную процедуру последовательного применения расширенного теста Дики-Фуллера (ADF), сопровождая ее анализом коррелограмм АКФ и ЧАКФ как исходного ряда, так и его последовательных разностей первого и второго порядка.. Определить качество применения тестов на основе различных статистических характеристик, определить целесообразность включения лаговых переменных на основе информационных критериев. Определить возможную сезонную интегрируемость процесса, применив тест Дики-Фуллера-Хаза. В качестве информационных средств выполнения задания рекомендуется использовать R.

Результатом выполнения кейс-задания является отчёт по кейсу 2. К отчёту предъявляются следующие требования:

1. Чёткое формулирование поставленной цели исследования (*например: цель: Определить к какому типу случайных процессов относится временной ряд «_____» для возможности дальнейшего его моделирования с помощью инструмента анализа временных рядов*).
2. Формулирование задач, решение которых необходимо для достижения поставленной цели.

Например:

Задачи:

1. *Описать исходные данные, указать источник информации, указать период ретроспективы, обосновать необходимость построения прогноза для данного показателя. Указать метод восполнения пропущенных данных. Если показатель индекс, то указать метод его расчета. Проанализировать график временного ряда (скриншот), сделав предположение о наличии детерминированного тренда и сезонности.*
2. *На основе анализа коррелограмм АКФ и ЧАКФ сделать предположение о типе процесса, представленного временным рядом. Коррелограммы анализируются как для исходного ряда, так и для ряда первых и вторых разностей.*
3. *С помощью расширенного теста Дики-Фуллера провести идентификацию случайного процесса, применяя процедуру Доладо-Дженкинсона-Сосвила-Риверо*
4. *Сделать заключение о типе процесса.*

3. Описание в виде пунктов, тех действий, которые требуются для решения поставленных задач. Все рисунки и таблицы последовательно нумеруются и описываются. Каждый пункт решения поставленных задач сопровождается анализом принятого решения. При проведении статистического теста Дики-Фуллера обязательно выписывается нулевая и альтернативная гипотеза, а также уравнение, обосновывающее тест, формулируется принятие решения на обосновано выбранном уровне значимости, указывается критическая область отказа от нулевой гипотезы в пользу альтернативной. Подбор длины максимального лага запаздывания для включения лаговых переменных в тестируемые уравнения (тест Дики-Фуллера) производится исходя из минимума информационных критериев.

4. В заключении выписывается тестируемое уравнение для процесса, к которому относится исследуемый временной ряд.

Кейс-задание 3 (по модулю 2). Определение типов случайных процессов, представленные временными рядами, с учётом структурных изменений

Проанализировать ряд на наличие структурных изломов и скачков с помощью теста Кванда-Эндрюса. Ввести фиктивные переменные (ds , $ds1$, dt), отвечающие за тип структурного скачка/излома. В зависимости от типа процесса построить модели с включением соответствующих фиктивных переменных, отвечающих за тип структурного скачка, используя тест Перрона. Для DS – ds , $ds1$, Для TS – dt , $ds1$. Проверить статическую значимость коэффициентов при фиктивных переменных с помощью теста Стьюдента, а также общую значимость модели тестируемого уравнения. Сделать вывод о качестве построенных уравнений с учётом введения в модель данных о структурных скачках. Сделать вывод о типе процесса с учетом структурных изменений TS, DS ($I(1)$, $I(2)$ или $I(0)$), TS+DS.

Определить качество применения тестов на основе различных статистических характеристик, определить целесообразность включения лаговых переменных на основе информационных критериев. В качестве информационных средств выполнения задания рекомендуется использовать R.

Результатом выполнения кейс-задания является отчёт по кейсу 3. К отчёту предъявляются следующие требования:

1. Чёткое формулирование поставленной цели исследования (например: цель: Уточнить к какому типу случайных процессов относится временной ряд « _____ » с учётом наличия структурных изменений в динамике ряда для возможности дальнейшего его моделирования с помощью инструмента анализа временных рядов).
2. Формулирование задач, решение которых необходимо для достижения поставленной цели. Например: Задачи:
*Проанализировать график временного ряда (скриншот), сделав предположение о наличии структурных изменений в динамике ряда. Подтвердить наличие точек излома/скачков в тесте Квандта-Эндрюса.
С помощью расширенного теста Перрона провести идентификацию случайного процесса, проверяя наличие структурных скачков, «подскоков» или изломов.
Сделать заключение о типе процесса с учётом структурных изменений.*
3. Описание в виде пунктов, тех действий, которые требуются для решения поставленных задач. Все рисунки и таблицы последовательно нумеруются и описываются. Каждый пункт решения поставленных задач сопровождается анализом принятого решения. При

проведении статистического теста Перрона и Квандта-Эндрюса, обязательно выписывается нулевая и альтернативная гипотеза, а также уравнение, обосновывающее тест, формулируется принятие решения на обосновано выбранном уровне значимости, указывается критическая область отказа от нулевой гипотезы в пользу альтернативной. Подбор длины максимального лага запаздывания для включения лаговых переменных в тестируемые уравнения производится исходя из минимума информационных критериев.

4. В заключении выписывается тип процесса с указанием типа структурного изменения.

Кейс-задание 4 (по модулю 3). Построение моделей нестационарных случайных процессов ARIMA.

В качестве исходной информации выбрать тот же временной ряд, что и для кейс-задания 2-3. Исходя из ранее определенного типа процесса, перейти к стационарности. Идентифицировать порядки ARIMA- модели p и q на основе анализа коррелограмм и используя критерий Лjungа-Бокса. Оценить параметры ARIMA моделей методом наименьших квадратов. Провести графический анализ единичных корней построенной модели. Оценить качество модели на основе индекса детерминации R^2 . Провести селекцию моделей (отбор лучшей) либо на основе критериев Акайке и Шварца, либо на основе минимума суммы квадратов отклонений. Провести процедуру адекватности построенной модели на основе соответствия остатков (ошибок) модели процессу белого шума (на основе анализа коррелограмм), проверить также нормальность распределения остатков. В качестве информационных средств выполнения задания рекомендуется использовать R.

Повышенный уровень: Построение ARIMA/SARIMA с учетом структурных изломов и скачков (введение в модель фиктивных переменных, отвечающих за скачок в тенденции и/или излом).

Результатом выполнения кейс-задания является отчет по лабораторной кейсу 4. К отчету предъявляются следующие требования:

1. Четкое формулирование поставленной цели исследования (*например: цель: Прогнозирование показателя «_____» на основе построения модели ARIMA*).
2. Формулирование задач, решение которых необходимо для достижения поставленной цели.

Например:

Задачи:

1. перевести временной ряд измеряющий показатель _____ в стационарный в зависимости от того, к какому типу процессов относится моделируемый ряд;
2. на основе анализа коррелограмм АКФ и ЧАКФ, полученного стационарного ряда провести идентификацию порядков моделей ARIMA;
3. методом наименьших квадратов найти оценки параметров моделей ARIMA;
4. при построении нескольких статистически значимых моделей ARIMA провести процедуру их селекции, исходя из минимума информационных критериев;
5. проведение процедуры адекватности, отобранной модели; проверить условие обратимости модели.

3. Описание в виде пунктов, тех действий, которые требуются для решения поставленных задач. Все рисунки и таблицы последовательно нумеруются и описываются. Каждый пункт решения поставленных задач сопровождается анализом принятого решения. При проведении статистических тестов, обязательно выписывается нулевая и альтернативная гипотеза, формулируется принятие решения на обосновано выбранном уровне значимости, указывается критическая область отказа от нулевой гипотезы в пользу альтернативной.

4. В заключении выписывается отобранная адекватная модель с оцененными коэффициентами с указанием под оценками коэффициентов значений t-статистик в скобках или стандартных ошибок коэффициентов.

Например:

$$y_t = 1,2 + 0,7y_{t-1} - 0,2y_{t-2} - 0,3\varepsilon_{t-1} - 0,7\varepsilon_{t-2} + \varepsilon_t$$

(2,26) (3,41) (-4,53) (-3,12) (-2,94)

Кейс-задание 5 (по модулю 3). Построение моделей нестационарных случайных процессов SARIMA.

В качестве исходной информации выбрать тот же временной ряд, что и для кейс-задания 2-3. Исходя из ранее определенного типа процесса, перейти к стационарности. Определить способ вхождения сезонности в структуру ряда (мультипликативно или аддитивно). Для построения модели SARIMA определить лаг сезонного сдвига S (на основе анализа коррелограмм АКФ и ЧАКФ) и идентифицировать порядки P (сезонной авторегрессии) и Q (сезонной скользящей средней) на основе анализа коррелограмм. Оценить параметры SARIMA модели методом наименьших квадратов. Провести графический анализ единичных корней построенной модели. Оценить качество модели на основе индекса детерминации R². Провести селекцию моделей (отбор лучшей) либо на основе критериев Акайке и Шварца, либо на основе минимума суммы квадратов отклонений. Провести процедуру адекватности построенной модели на основе соответствия остатков (ошибок) модели процессу гауссова белого шума (на основе анализа коррелограмм и теста Шапиро-Уилка). В качестве информационных средств выполнения задания рекомендуется использовать R.

Повышенный уровень: Построение SARIMA с учетом структурных изломов и скачков (введение в модель фиктивных переменных, отвечающих за скачок в тенденции и/или излом).

Результатом выполнения кейс-задания является отчет по кейсу 5. К отчету предъявляются следующие требования:

1. Чёткое формулирование поставленной цели исследования (*например: цель: Прогнозирование показателя «_____» на основе построения модели SARIMA*).
2. Формулирование задач, решение которых необходимо для достижения поставленной цели.

Например:

Задачи:

1. перевести временной ряд измеряющий показатель «_____» в стационарный в зависимости от того, к какому типу процессов относится моделируемый ряд (в том числе с учётом сезонной интегрируемости процесса);
2. на основе анализа коррелограмм АКФ и ЧАКФ, полученного стационарного ряда провести идентификацию порядков моделей SARIMA; определить каким образом (мультипликативно или аддитивно) входит сезонная компонента в структуру ряда.
3. методом наименьших квадратов найти оценки параметров моделей SARIMA;
4. при построении нескольких статистически значимых моделей SARIMA провести процедуру их селекции, исходя из минимума информационных критериев;
5. проведение процедуры адекватности, отобранной модели;

3. Описание в виде пунктов, тех действий, которые требуются для решения поставленных задач. Все рисунки и таблицы последовательно нумеруются и описываются. Каждый пункт решения поставленных задач сопровождается анализом принятого решения. При проведении статистических тестов, обязательно выписывается нулевая и альтернативная гипотеза, формулируется принятие решения на обосновано выбранном уровне значимости, указывается критическая область отказа от нулевой гипотезы в пользу альтернативной.
4. В заключении выписывается отобранная адекватная модель с оценёнными коэффициентами с указанием под оценками коэффициентов значений t-статистик в скобках или стандартных ошибок коэффициентов.

Например: для периода сезонности $T=12$:

$$y_t = 1,2 + 0,7y_{t-12} - 0,2y_{t-24} - 0,3\varepsilon_{t-12} - 0,7\varepsilon_{t-24} + \varepsilon_t$$

(2,26) (3,41) (-4,53) (-3,12) (-2,94)

Кейс-задание 6 (по модулю 3). Построение моделей нестационарных случайных процессов ARIMAX с экзогенными переменными.

В качестве исходной информации выбрать тот же временной ряд, что и для кейс-задания 2-3, в качестве экзогенной переменной выбрать процесс той же длины, оказывающий теоретически влияние на целевую моделируемую переменную. Определить тип процесса экзогенной переменной. Исходя из ранее определенного типа процесса целевой (зависимой) и экзогенной переменных, перейти к стационарности. Идентифицировать порядки ARIMA- модели p и q на основе анализа коррелограмм. Проанализировать кросскоррелограммы между целевой (зависимой) и экзогенной переменными, провести тест Гренджера на причинность. Оценить параметры ARIMAX модели методом наименьших квадратов. Провести графический анализ единичных корней построенной модели. Оценить качество модели на основе индекса детерминации R^2 . Провести селекцию моделей (отбор лучшей) либо на основе критериев Акайке и Шварца, либо на основе минимума суммы квадратов отклонений. Провести процедуру адекватности построенной модели на основе соответствия остатков (ошибок) модели процессу гауссова белого шума (на основе анализа коррелограмм и теста Шапиро-Уилка). В качестве информационных средств выполнения задания рекомендуется использовать R.

Результатом выполнения кейс-задания является отчёт по кейсу 6. К отчёту предъявляются следующие требования:

5. Чёткое формулирование поставленной цели исследования (например: *цель: Прогнозирование показателя « _____ » на основе построения модели ARIMAX*).
6. Формулирование задач, решение которых необходимо для достижения поставленной цели.

Например:

Задачи:

6. перевести временной ряд измеряющий показатель « _____ » в стационарный в зависимости от того, к какому типу процессов относится моделируемый ряд;
7. перевести временной ряд измеряющий показатель « _____ », в качестве экзогенной переменной, в стационарный в зависимости от того, к какому типу процессов относится моделируемый ряд;

8. на основе анализа коррелограмм АКФ и ЧАКФ, а также кросс-коррелограмм и теста на ричинность Гренджера, полученного стационарного ряда провести идентификацию порядков моделей ARIMAX;

9. методом наименьших квадратов найти оценки параметров моделей ARIMAX;

10. при построении нескольких статистически значимых моделей ARIMAX провести процедуру их селекции, исходя из минимума информационных критериев;

11. проведение процедуры адекватности, отобранной модели;

7. Описание в виде пунктов, тех действий, которые требуются для решения поставленных задач. Все рисунки и таблицы последовательно нумеруются и описываются. Каждый пункт решения поставленных задач сопровождается анализом принятого решения. При проведении статистических тестов обязательно выписывается нулевая и альтернативная гипотеза, формулируется принятие решения на обосновано выбранном уровне значимости, указывается критическая область отказа от нулевой гипотезы в пользу альтернативной.

8. В заключении выписывается отобранная адекватная модель с оценёнными коэффициентами с указанием под оценками коэффициентов значений t-статистик в скобках или стандартных ошибок коэффициентов.

Например, если x_t – является экзогенной переменной, то:

$$y_t = 1,2 + 0,7y_{t-1} - 0,2y_{t-2} - 0,3\varepsilon_{t-1} - 0,7\varepsilon_{t-2} + 0,6x_t + \varepsilon_t$$

(2,26) (3,41) (-4,53) (-3,12) (-2,94) (3,8)

Кейс-задача 7 (модуль 4) Адаптивные методы краткосрочного прогнозирования

1. Рассчитать параметры адаптивных моделей (Хольта, Хольта-Уинтерса) с помощью R Studio. В трех режимах задание параметров адаптации: вручную, по сетке поиска (режим Grid search) и автоматический (Скриншорты всех анализируемых таблиц (кроме таблиц данных) и графиков приложить). Выбрать лучшие параметры адаптации (выбор обосновать).

2. Проверить адекватности построенной модели с помощью анализа остатков на соответствие гауссову белому шуму (на основе коррелограмм и теста Шапиро-Уилка).

В качестве информационных средств выполнения задания рекомендуется использовать R.

Результатом выполнения задания является отчёт кейсу 7. К отчёту предъявляются следующие требования:

1. Чёткое формулирование поставленной цели исследования (например: *цель: Прогнозирование показателя « _____ на основе построения модели адаптивного сглаживания*).
2. Формулирование задач, решение которых необходимо для достижения поставленной цели. Формулируется заключение о качестве построенной модели.

Кейс-задача 8 (модуль 5) Обобщённое прогнозирование:

Построение обобщённого прогнозирования: построить прогноз по четырём моделям ARIMA, ARIMAX, SARIMA и адаптивных моделей, провести содержательный анализ прогнозов, согласно построенным моделям. Прогноз строится на период упреждения $T/4$, где T – общее число временных интервалов исходного ряда. Применить технологию валидации, приемлемую для анализа временных рядов.

Оценить метрики качества для статического и динамического прогноза по всем моделям, дать интерпретацию прогнозных свойств моделей.

Рассчитать показатели:

- Квадратный корень средней ошибки предсказания;
- Среднюю ошибку прогноза по модулю
- Среднюю процентную ошибку по модулю (среднюю ошибку аппроксимации)
- Коэффициент неравенства Тейла;
- Долю систематической ошибки;
- Долю вариации;
- Долю ковариации.

Сделать вывод о выборе лучшей модели прогнозирования.

Результатом выполнения кейса является отчёт, содержащий графики исходных данных и прогнозов (в трех вариантах: базовый, пессимистический и оптимистический с учётом 2-х стандартных отклонений при уровне доверительной вероятности 0,95) для всех четырёх моделях, а также метрики прогнозного качества моделей, и заключение о приемлемости модели прогнозирования на их основе.

8.5.

Наименование модуля	Задание	Балл	Критерии оценки
Входное тестирование	ТЕСТ	10	Проходной балл -5
Модуль 1 – Введение в курс	Кейс 1	10	Система R должна быть загружена (минимально допустимый балл - 8 баллов)
Модуль 2. Типы случайных процессов, представленные временными рядами	Кейс 2	10	Для минимального балла (5) должны быть выполнены все задания, но могут быть допущены неточности в интерпретации результатов анализа данных
	Кейс 3	10	Для минимального балла (5) должны быть выполнены все задания, но могут быть допущены неточности в интерпретации результатов анализа данных
	Тест	8	Проходной балл - 4
Модуль 3. Модели нестационарных случайных процессов ARIMA/SARIMA/ARIMAX	Кейс 4	10	Для минимального балла (5) должны быть выполнены все задания, но могут быть допущены неточности в интерпретации результатов моделирования
	Кейс 5	10	Для минимального балла (5) должны быть выполнены все задания, но могут быть допущены неточности в интерпретации результатов моделирования
	Кейс 6	12	Для минимального балла (5) должны быть выполнены все задания, но могут быть допущены неточности в интерпретации результатов моделирования
	Тест	10	Проходной балл - 6
Модуль 4. Модели адаптивного краткосрочного прогнозирования	Кейс 7	10	Для минимального балла (5) должны быть выполнены все задания, но могут быть допущены неточности в интерпретации результатов моделирования
	Тест	6	Проходной балл - 3

Модуль 5 Обобщённое прогнозирование	Кейс 8	8	Для минимального балла (4) должны быть выполнены все задания, но могут быть допущены неточности в интерпретации результатов оценки качества прогноза
Итоговая аттестация (итоговый проект)	Задание на проект	40	Итоговое задание считается выполненным на минимально допустимый балл 20, если есть полностью сформированное задание, но отсутствует интерпретация полученных результатов
		Минимальный балл для получения зачета по КПК - 80, максимальный - 154	

9. Организационно-педагогические условия реализации программы

9.1. Кадровое обеспечение программы

№ п/п	Фамилия, имя, отчество (при наличии)	Место основной работы и должность, ученая степень и ученое звание (при наличии)	Ссылки на веб-страницы с портфолио (при наличии)	Фото в формате jpeg	Отметка о полученном согласии на обработку персональных данных
1	Лакман Ирина Александровна	ФГБОУ ВО «Башкирский государственный университет», заведующая лабораторией исследования социально-экономических проблем регионов, к.т.н., доцент		Загружено на платформу	Да

9.2. Учебно-методическое обеспечение и информационное сопровождение

Учебно-методические материалы	
Методы, формы и технологии	Методические разработки, материалы курса, учебная литература
Методы организации учебно-познавательной деятельности: практический; Форма: дистанционная; Технологии: Информационно – коммуникационная технология; Кейс технология	<ol style="list-style-type: none"> 1. Анализ данных : учебник для академического бакалавриата / ГУ - Высшая школа экономики; под ред. В. С. Мхитаряна .— Москва : Юрайт, 2016 .— 490 с. (13 экз в библиотеке) 2. Ананьев, В. А. Анализ экспериментальных данных [Электронный ресурс] : учеб. пособие / В. А. Ананьев .— Кемерово : Кемеровский государственный университет, 2009 .— 102 с. [Электронный ресурс] URL=https://biblioclub.ru/index.php?page=book_red&id=232208 3. Эконометрика : учебник для магистров / Санкт-Петербургский государственный экономический университет ; под ред. И. И. Елисеевой .— Москва : Юрайт, 2014 .— 449 с. : ил. ; 21 см .— (Магистр) .— ОГЛАВЛЕНИЕ кликните на URL-> .— Библиогр.: с. 430-432 .—

	<p><URL:http://www.library.ugatu.ac.ru/pdf/teach/Ekonometrika_Eliseeva_2014.pdf >.</p> <p style="text-align: center;">Дополнительная литература</p> <p>1. Эконометрика : учебник / под ред. В. С. Мхитаряна .— Москва : Проспект, 2014 .— 384 с. :</p> <p>2. Канторович, Г. (канд. физ.-мат. наук, проф., проректор) . Роберт Энгл и Клайв Гренджер: новые области экономических исследований : Нобелевская премия 2003 года по экономике / Г. Канторович, М. Турунцева // Вопросы экономики .— 2004 .— N 1 .— С.37-48.</p> <p>3. Федорова, Е. А. (кандидат экономических наук ; доцент) . Финансовая интеграция фондовых рынков стран БРИК: эконометрический анализ [Текст] / Е. А. Федорова // Финансы и кредит .— 2011 .— N 18 .— С. 24-29 : табл. — (Фондовый рынок) .</p>

Информационное сопровождение	
Электронные образовательные ресурсы	Электронные информационные ресурсы
http://sdo.bashedu.ru/course/view.php?id=3107	https://www.kaggle.com/

9.3. Материально-технические условия реализации программы

Вид занятий	Наименование оборудования, программного обеспечения
Лекции, практические занятия	<p>Аппаратные требования Intel Pentium или аналогичный процессор с тактовой частотой 300MHz и выше. SVGA монитор, с разрешением экрана, как минимум, 800x600 точек и глубиной цвета 16 bit (рекомендуемое разрешение экрана — 1024x768). Звуковая карта, акустическая система или наушники. Доступ в Интернет со скоростью 56 кбит/с и выше.</p> <p>Программное обеспечение Операционная система: Windows 7 или более продвинутая, Macintosh, Linux Браузер: Internet Explorer 7 или более продвинутый, Mozilla Firefox (скачать бесплатно: http://www.mozilla.org/download.html) и т.п.</p> <p>Для просмотра электронных версий учебных курсов необходимо наличие установленных программ: Microsoft Internet Explorer 7.0 и выше (Загрузить с сайта www.microsoft.com)</p>

	Adobe Flash Player версии 7.0 и выше (Загрузить с сайта http://www.adobe.com/)
--	---