МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ

УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

(НИЯУ МИФИ)

Институт Финансовых Технологий и Экономической Безопасности

Кафедра Финансового Мониторинга

Лабораторная работа №5:

По курсу «Макростатический анализ и прогнозирование»

Работу выполнил: студент группы С18-712: Луканов А.В.

Проверила: Домашова Д. В.

Москва 2021

# 1. Постановка задачи

По данным о заболеваемости коронавирусом в Бельгии и Самарской области для исследуемого временного ряда:

1) на основе визуального анализа, а также анализа автокорреляционной и частной автокорреляционной функций определить компонентный состав;

2) построить АРПСС-модель;

3) исследовать адекватность модели;

4) построить прогноз на 3 периода вперед.

# 2. Ход работы. Бельгия.

Исходные данные по заболеваемости (Рисунок 1).

|  |  |
| --- | --- |
| **date** | **new\_cases** |
| 2020-02-21 | 0 |
| 2020-02-22 | 0 |
| 2020-02-23 | 0 |
| 2020-02-24 | 0 |
| 2020-02-25 | 0 |
| 2020-02-26 | 0 |
| 2020-02-27 | 0 |
| 2020-02-28 | 0 |
| 2020-02-29 | 0 |
| 2020-03-01 | 1 |
| 2020-03-02 | 6 |
| 2020-03-03 | 5 |
| 2020-03-04 | 10 |
| 2020-03-05 | 27 |
| 2020-03-06 | 59 |
| 2020-03-07 | 60 |
| 2020-03-08 | 31 |
| 2020-03-09 | 39 |
| 2020-03-10 | 28 |
| 2020-03-11 | 47 |
| 2020-03-12 | 0 |
| 2020-03-13 | 245 |
| 2020-03-14 | 130 |
| 2020-03-15 | 197 |

Рисунок 1 – Исходные данные по Бельгии

Был построен график временного ряда (Рисунок 2).

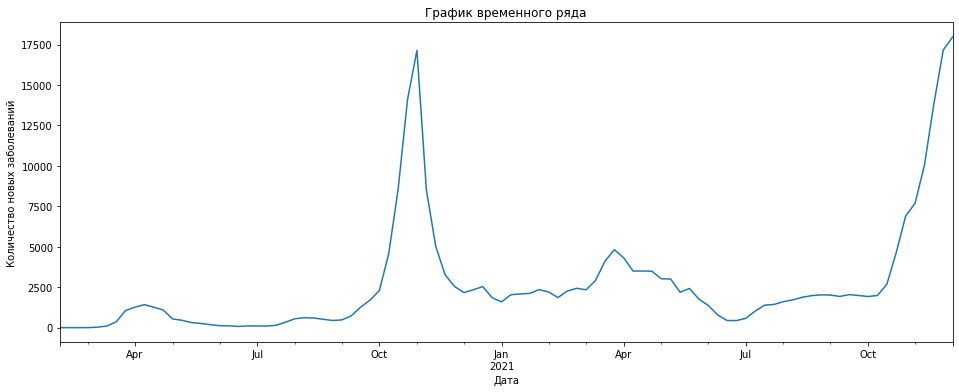


Рисунок 2 – График временного ряда

Определим тип ряда. Для этого сначала оценим модель с константой и линейным трендом (Рисунок 3).

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 3 – Определение типа ряда, шаг 1.

Критерий Дики-Фуллера говорит о том, что ряд нестационарен, коэффициент при тренде не значим. Построим модель авторегрессии 1-го порядка с константой (Рисунок 4).

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 4 – Определение типа ряда, шаг 2.

Критерий Дики-Фуллера говорит о том, что ряд нестационарен, коэффициент при константе незначимый. Построим модель авторегрессии 1-го порядка без константы (Рисунок 5).

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 5 –Определение типа ряда, шаг 3

Критерий Дики-Фуллера говорит о том, что ряд нестационарен. Ряд можно отнести к классу DS(Случайное блуждание без дрейфа)

Устанавливаем порядок интеграции d, с помощью расширенного критерия Дики-Фуллера (Рисунок 6). Ряд нестационарен при d = 0.

Изображение выглядит как текст

Автоматически созданное описание

Рисунок 6 – Расширенный критерий Дики-Фуллера

Построим графики автокорреляционной и частной автокорреляционной функции для определения параметров p и q (Рисунок 7). Выберем p=2 (2 первых значимых лага на ЧАКФ) и q=3 (3 первых значимых лага на АКФ).



Рисунок 7 - Автокорреляционная и частная автокорреляционная функции

Переберем модели с различными значениями параметров p, d, q и выведем 5 с наименьшим значением информационного критерия Байеса (Рисунок 8).

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 8 – результат перебора параметров.

В результате перебора лучшими параметрами оказались p=0, d=2, q=2. Построим график вторых разностей ряда, а также его АКФ и ЧАКФ (Рисунок 9).

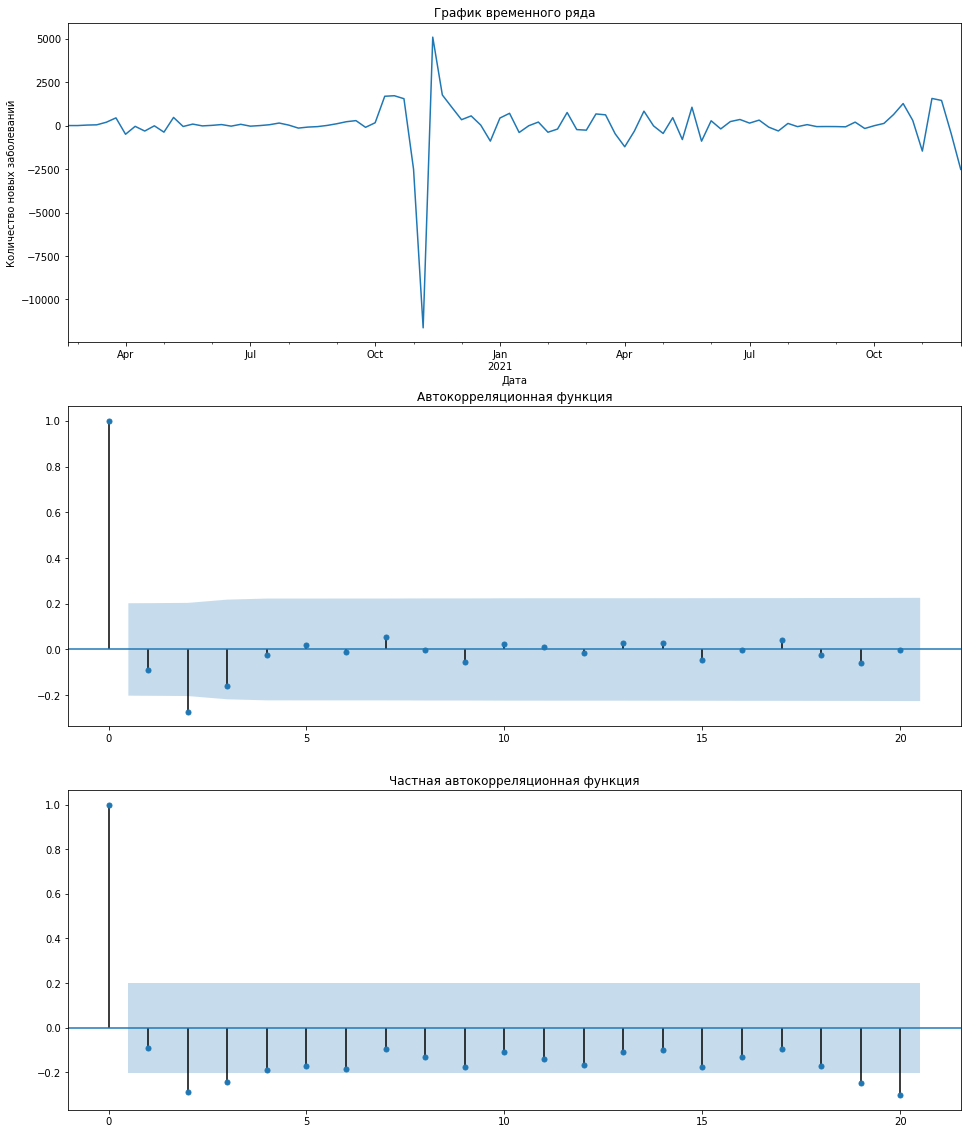


Рисунок 9 – график ряда двойных разностей, его АКФ и ЧАКФ

Оценим лучшую модель (Рисунок 10).

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 10 – оценка наилучшей модели.

Оценка модели АРПСС(0,2,2) выглядит следующим образом :

Построим график предсказанных и исходных значений (Рисунок 11) и рассчитаем функционалы качества модели (Рисунок 12).

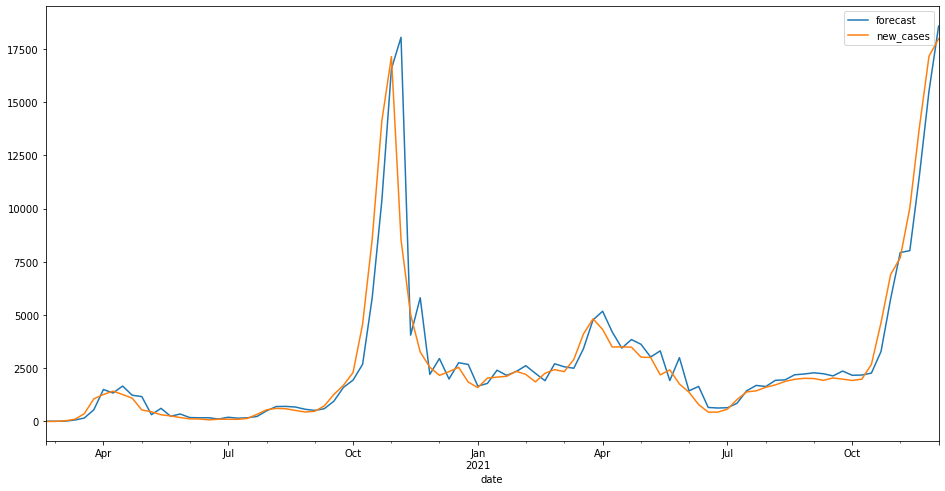


Рисунок 11 - график предсказанных и исходных значений

Изображение выглядит как текст

Автоматически созданное описание

Рисунок 11 – функционалы качества модели

Проверим модель на адекватность. Для этого проверим нормальность распределения остатков модели: проверим их с помощью критерия Дики-Фуллера (рисунок 12), построим их график, АКФ и ЧАКФ (рисунок 13), сравним график распределения остатков с нормальным (рисунок 14) и построим qq-график остатков (рисунок 15)

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 12 – критерий Дики-Фуллера для остатков

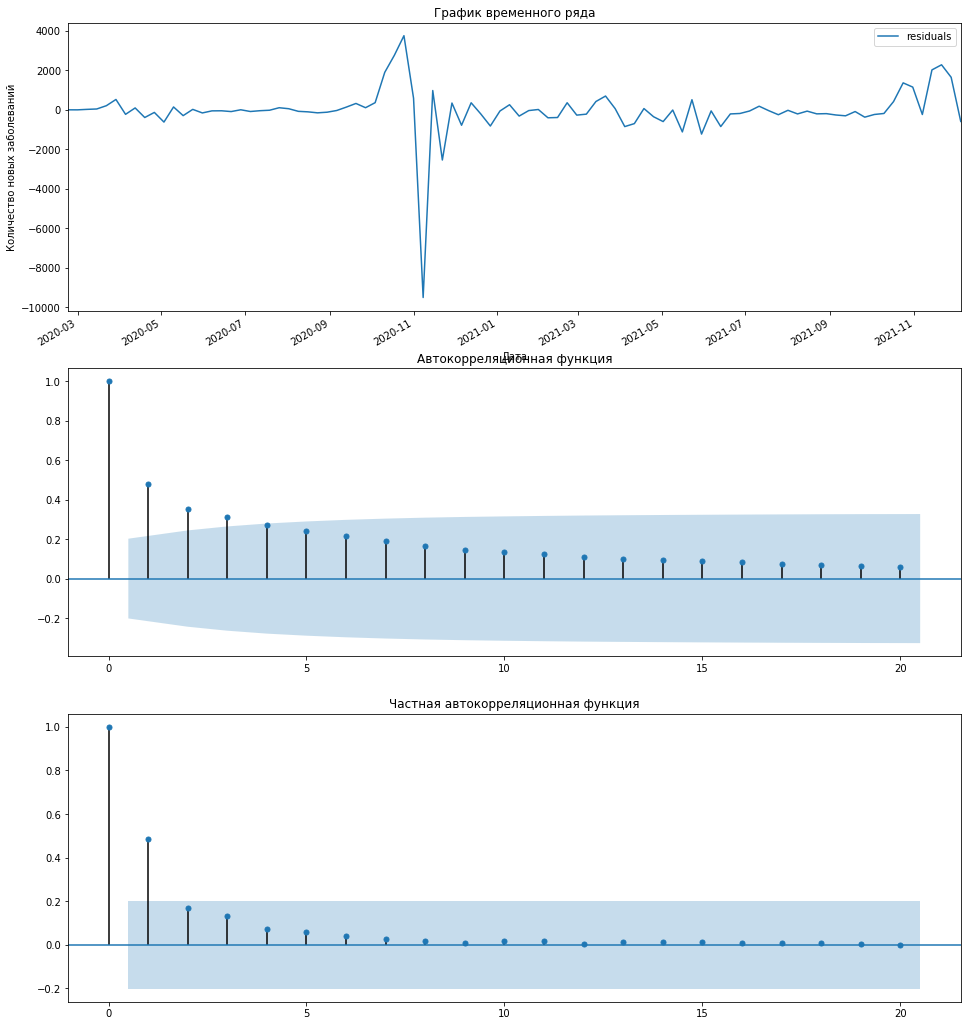


Рисунок 13 – График остатков, их АКФ и ЧАКФ

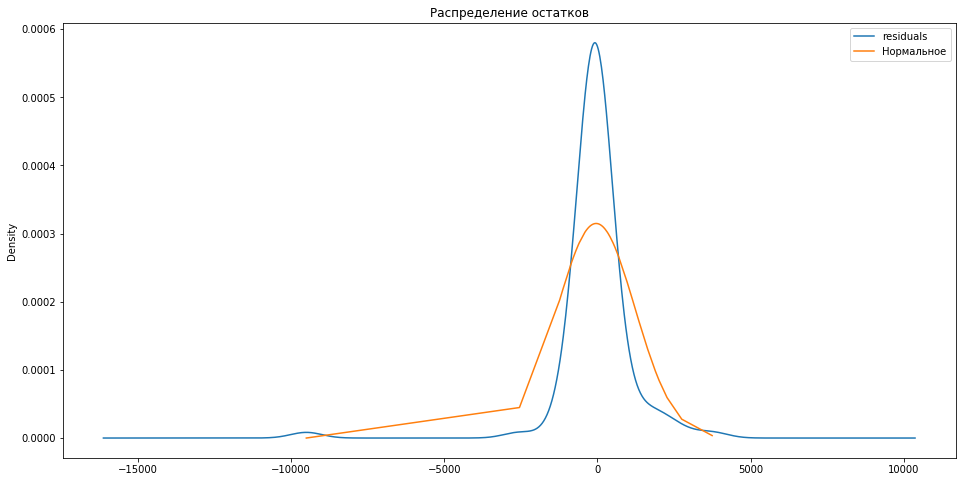


Рисунок 14 – Сравнение распределения остатков с нормальным

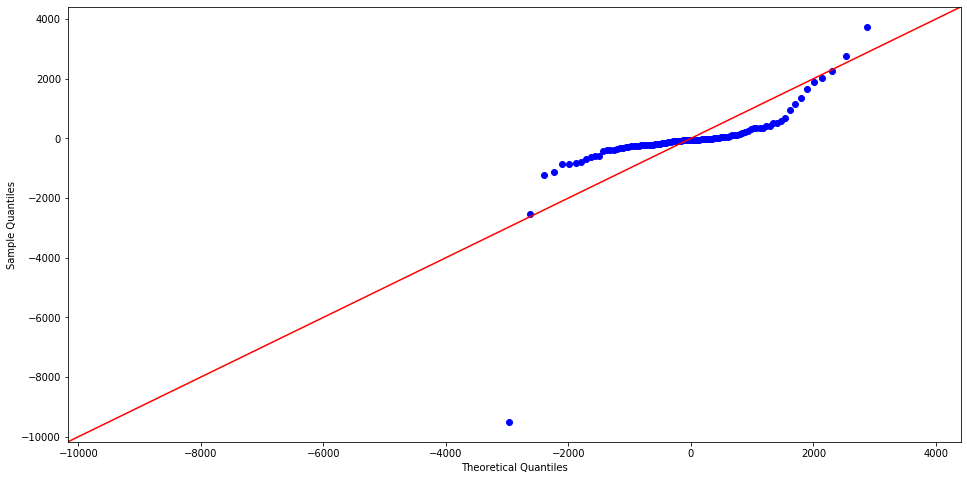


Рисунок 15 – qq-график остатков

Исходя из графиков АКФ и ЧАКФ требование к некоррелированности остатков не выполнено, кроме того, остатки распределены не нормально, следовательно модель не является адекватной.

Спрогнозируем будущие значения ряда на 2 недели вперед (Рисунок 16).

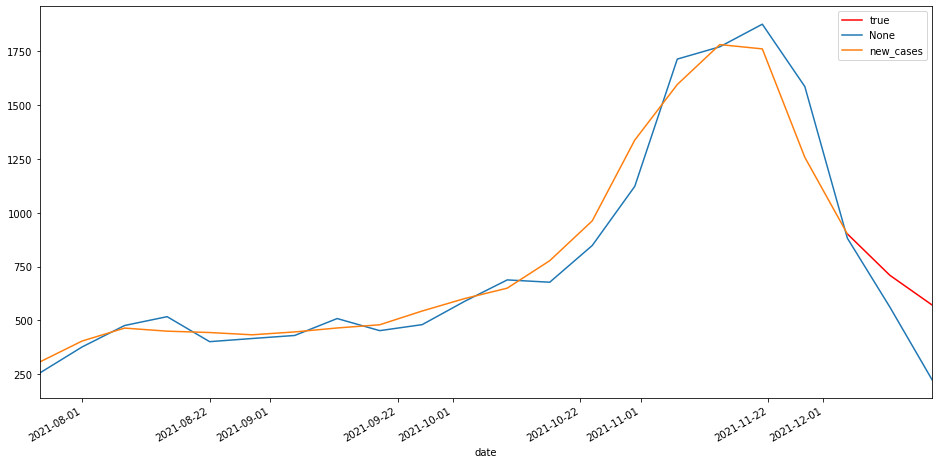


Рисунок 16 – прогноз на 2 недели

Добавим в модель сезонность и построим модель SARIMA. Возьмем сдвиг по сезонности равный 6 месяцам. Проверим модель с помощью критерия Дики-Фуллера (Рисунок 17). Ряд стационарен, следовательно допустимо взять порядок интегрированной сезонной разности D=0 и D=1.

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 17 – Критерий Дики-Фуллера для сезонной модели

Также были построены графики для сезонной модели (Рисунок 18), АКФ и ЧАКФ (Рисунок 19) для D=0.

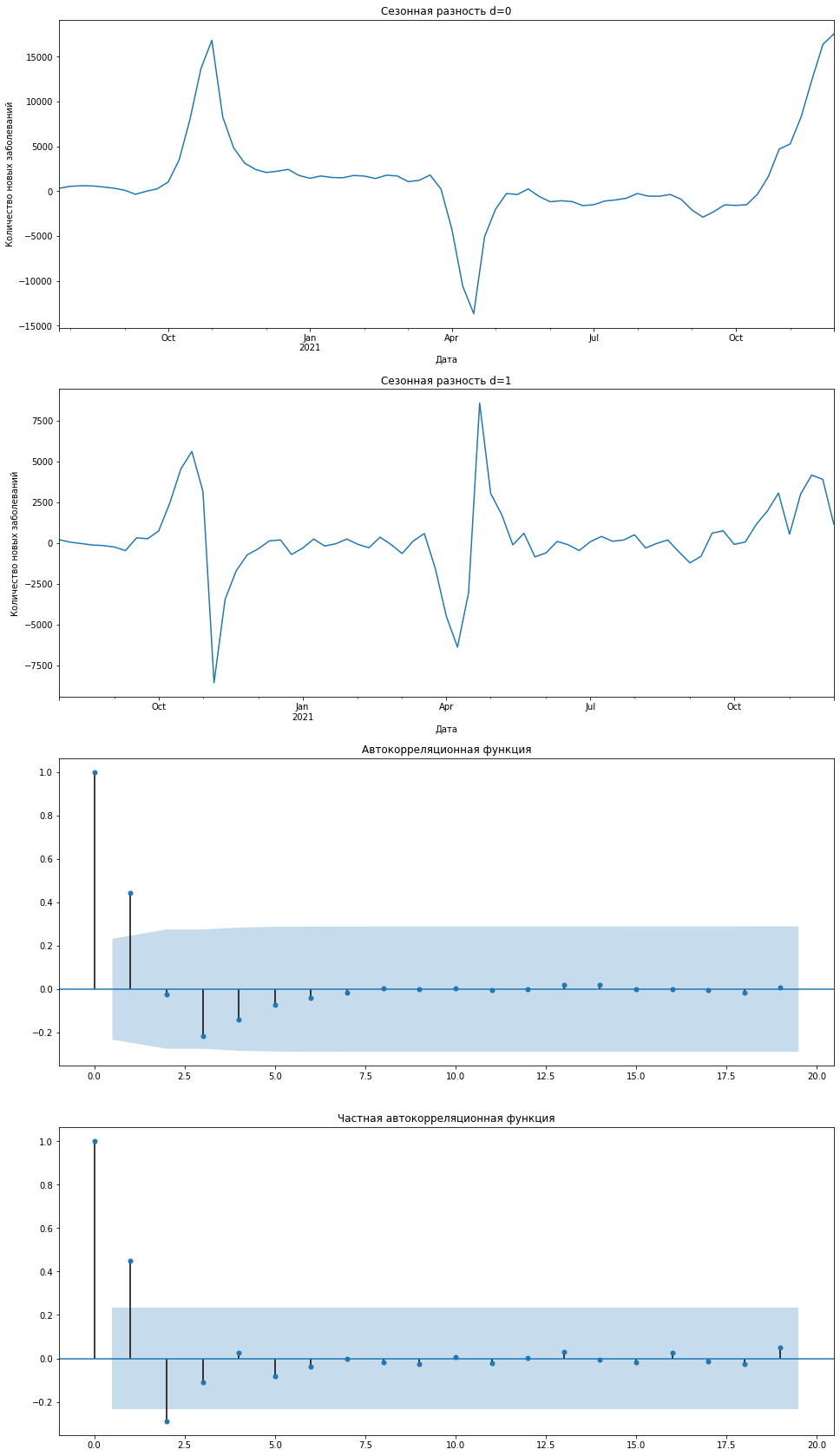


Рисунок 18 – Графики для сезонной модели

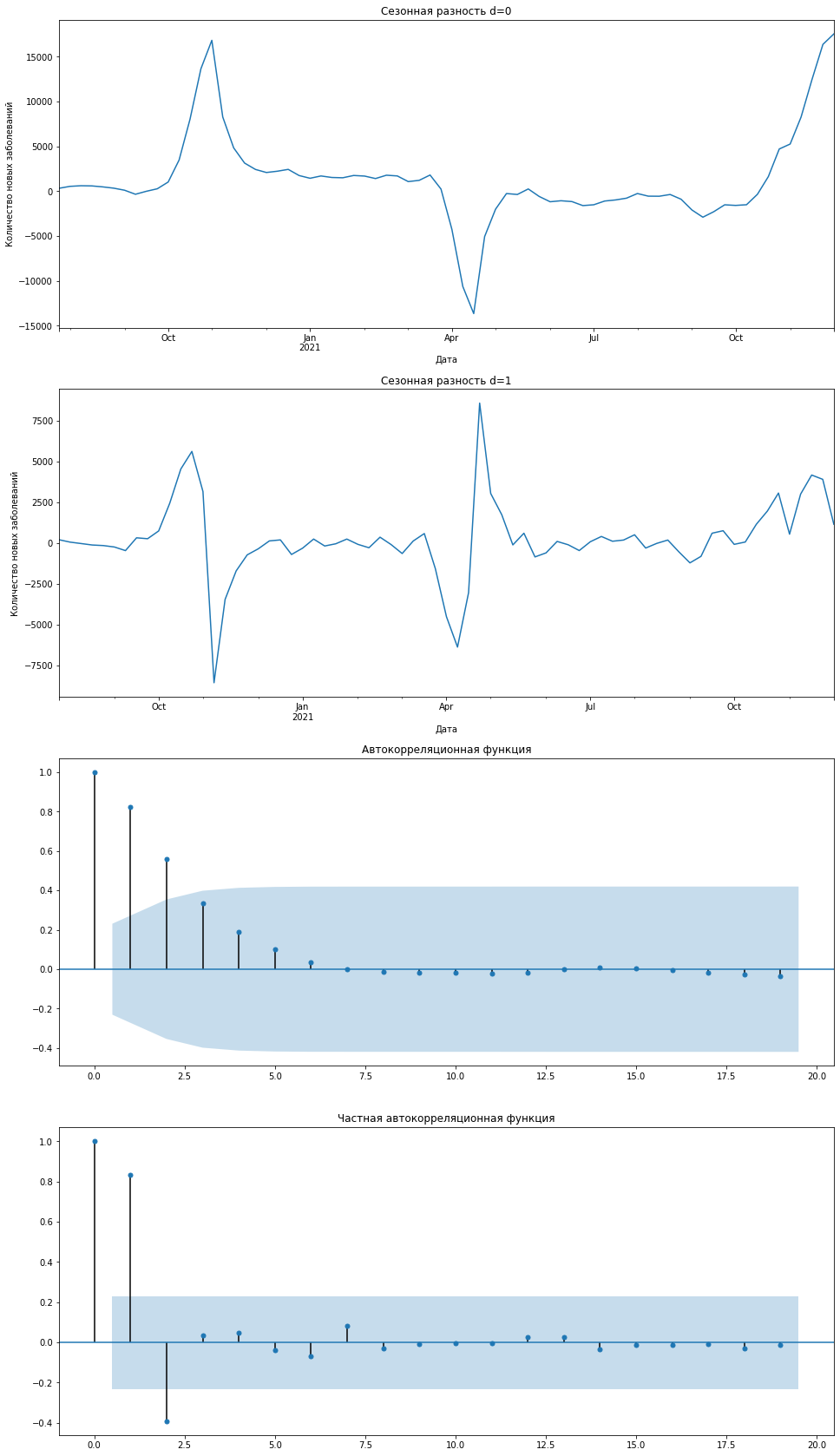


Рисунок 19 – АКФ и ЧАКФ сезонной модели

По графикам АКФ и ЧАКФ определяем P и Q. Выберем P = 2 (последний лаг кратный размеру сезонности на ЧАКФ - первый), Q = 2 (последний лаг кратный размеру сезонности на АКФ – первый). Используя лучшие значения p, d, q из предыдущей модели строим модель SARIMA (Рисунок 20).

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 20 – оценка модели SARIMA

Проведем анализ остатков. Основная гипотеза Льюнга-Бокса (об отсутствии автокорреляции) принимается. В тесте на гетероскедастичность гипотеза о гомоскедастичности остатков принимается. Согласно тесту Харке-Бера – остатки распределены не нормально. Построим график значений сезонной модели (Рисунок 21) и рассчитаем функционалы качества (Рисунок 22).

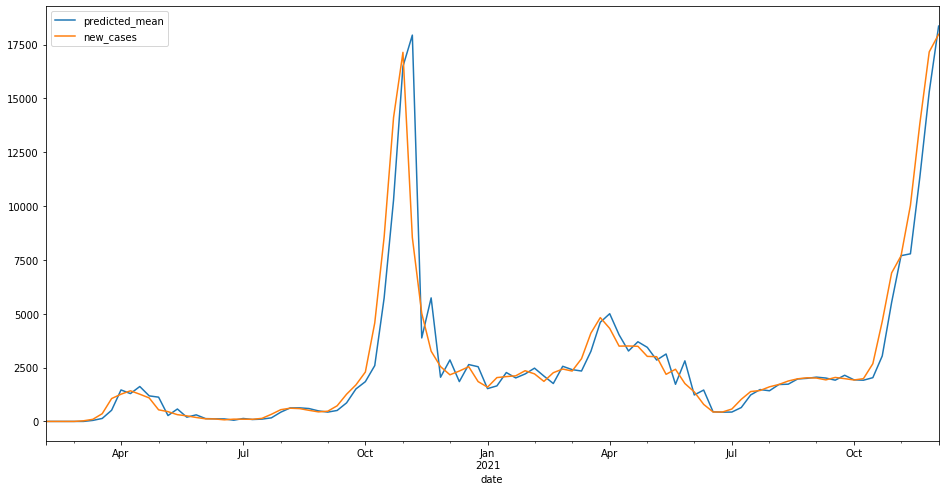


Рисунок 21 – график значений модели SARIMA

Изображение выглядит как текст

Автоматически созданное описание

Рисунок 22 – функционалы качества модели SARIMA

Проверим остатки модели на нормальность. Для этого проверим их с помощью критерия Дики-Фуллера (рисунок 23), построим их график, АКФ и ЧАКФ (рисунок 24), сравним график распределения остатков с нормальным (рисунок 25) и построим qq-график остатков (рисунок 26)

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 23 – критерий Дики-Фуллера для остатков

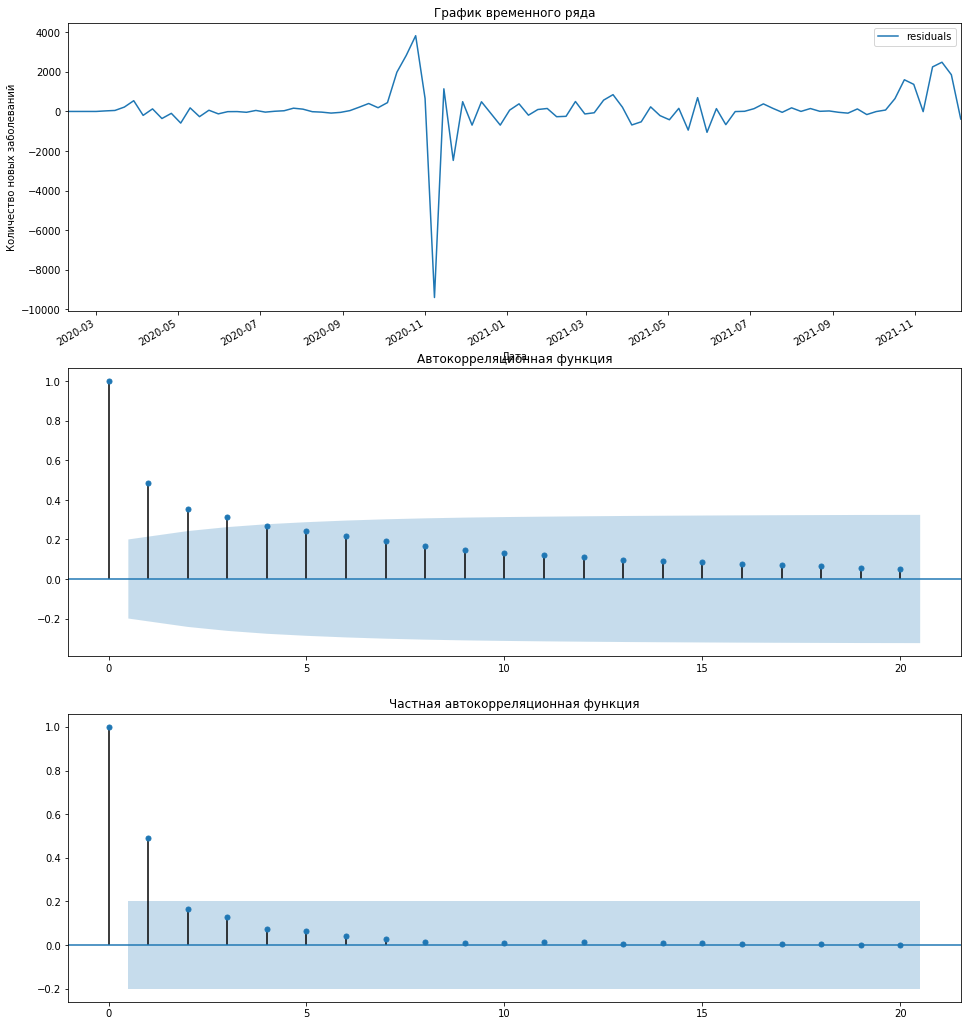


Рисунок 24 – График остатков, их АКФ и ЧАКФ

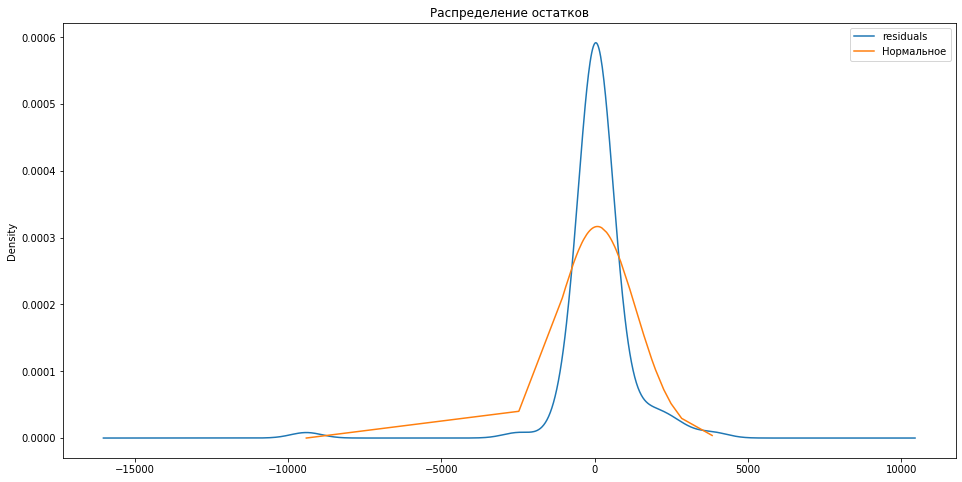


Рисунок 25 – Сравнение распределения остатков с нормальным

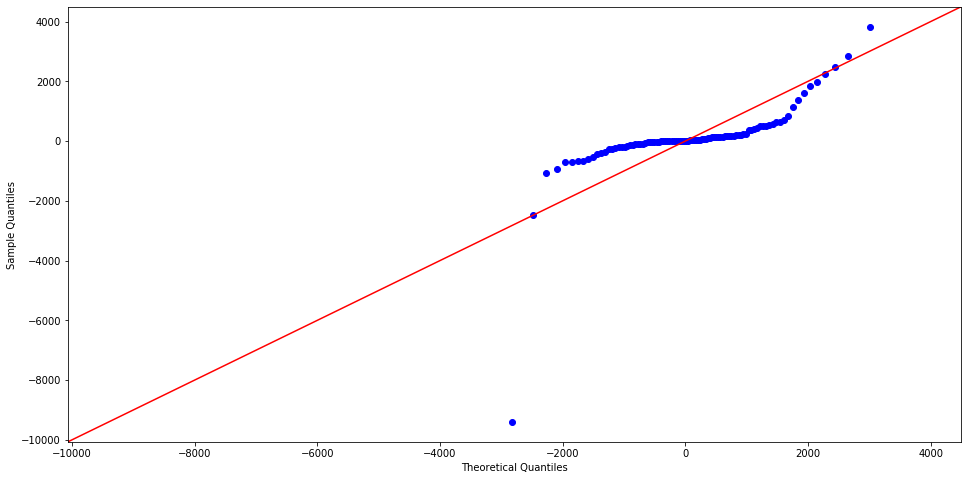


Рисунок 26 – qq-график остатков

Исходя из графиков остатки распределены не нормально, следовательно модель не является адекватной.

Спрогнозируем будущие значения ряда на 2 недели вперед (Рисунок 27).

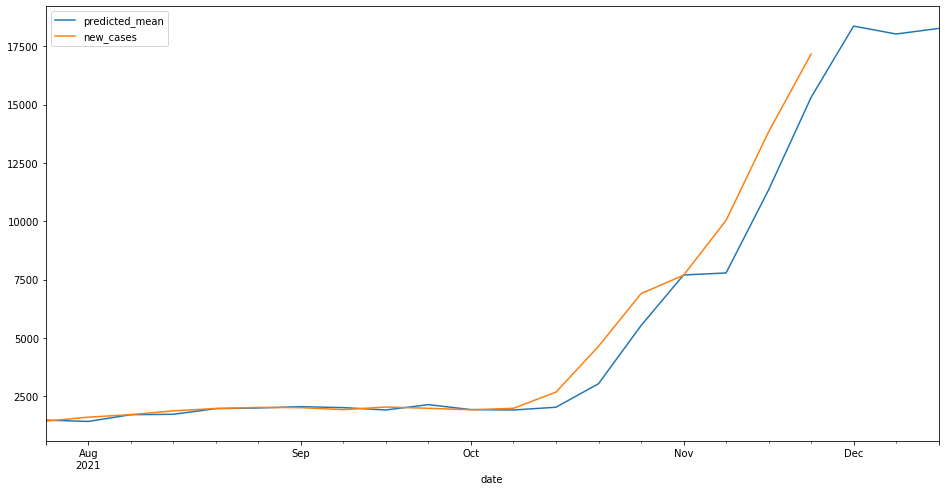


Рисунок 27 – прогноз на 2 недели

Рисунок 28 – график заболеваемости и смертности

# 3. Ход работы. Самарская область.

Исходные данные по заболеваемости (Рисунок 1).

|  |  |
| --- | --- |
| **date** | **new\_cases** |
| 19.03.2020 | 3 |
| 20.03.2020 | 2 |
| 21.03.2020 |  |
| 22.03.2020 |  |
| 23.03.2020 | 4 |
| 24.03.2020 |  |
| 25.03.2020 |  |
| 26.03.2020 |  |
| 27.03.2020 | 1 |
| 28.03.2020 |  |
| 29.03.2020 | 2 |
| 30.03.2020 | 2 |
| 31.03.2020 |  |
| 01.04.2020 |  |
| 02.04.2020 |  |
| 03.04.2020 | 1 |
| 04.04.2020 |  |
| 05.04.2020 | 2 |
| 06.04.2020 |  |
| 07.04.2020 |  |
| 08.04.2020 |  |
| 09.04.2020 |  |
| 10.04.2020 |  |
| 11.04.2020 | 1 |

Рисунок 1 – Исходные данные по Оренбургской области

Был построен график временного ряда (Рисунок 2).

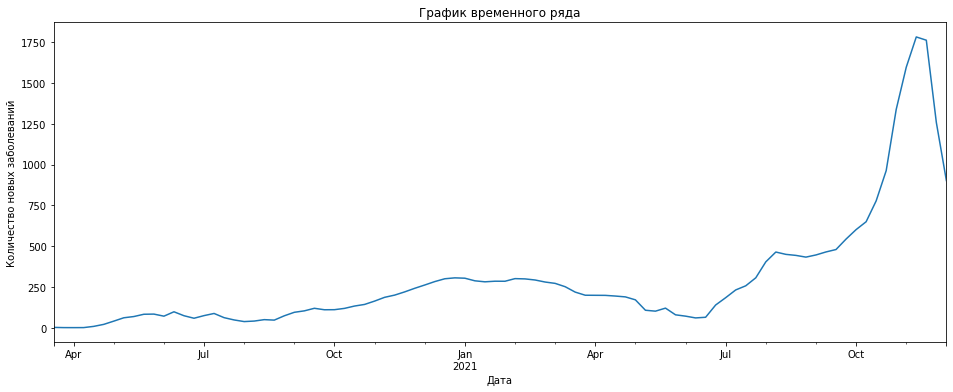


Рисунок 2 – График временного ряда

Определим тип ряда. Для этого сначала оценим модель с константой и линейным трендом (Рисунок 3).

Изображение выглядит как текст, квитанция

Автоматически созданное описание

Рисунок 3 – Определение типа ряда, шаг 1.

Критерий Дики-Фуллера говорит о том, что ряд нестационарен, поэтому строим модель только с константой и трендом. Коэффициент при тренде незначим, значит тренд отсутствует. Построим модель авторегрессии 1-го порядка с константой (Рисунок 4).

Изображение выглядит как текст, квитанция

Автоматически созданное описание

Рисунок 4 – Определение типа ряда, шаг 2.

Критерий Дики-Фуллера говорит о том, что ряд нестационарен, построим модель только с константой. Коэффициент при константе не значим, значит, процесс не является случайным блужданием с дрейфом. Построим модель авторегрессии 1-го порядка без константы (Рисунок 5).

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 4 – Определение типа ряда, шаг 3.

Критерий Дики-Фуллера говорит о том, что ряд нестационарен. Ряд можно отнести к классу DS(Случайное блуждание без дрейфа)

Устанавливаем порядок интеграции d, с помощью расширенного критерия Дики-Фуллера (Рисунок 5).

Изображение выглядит как текст

Автоматически созданное описание

Рисунок 5 – Расширенный критерий Дики-Фуллера

Построим графики вторых разностей и их автокорреляционной и частной автокорреляционной функции для определения параметров p и q (Рисунок 6).

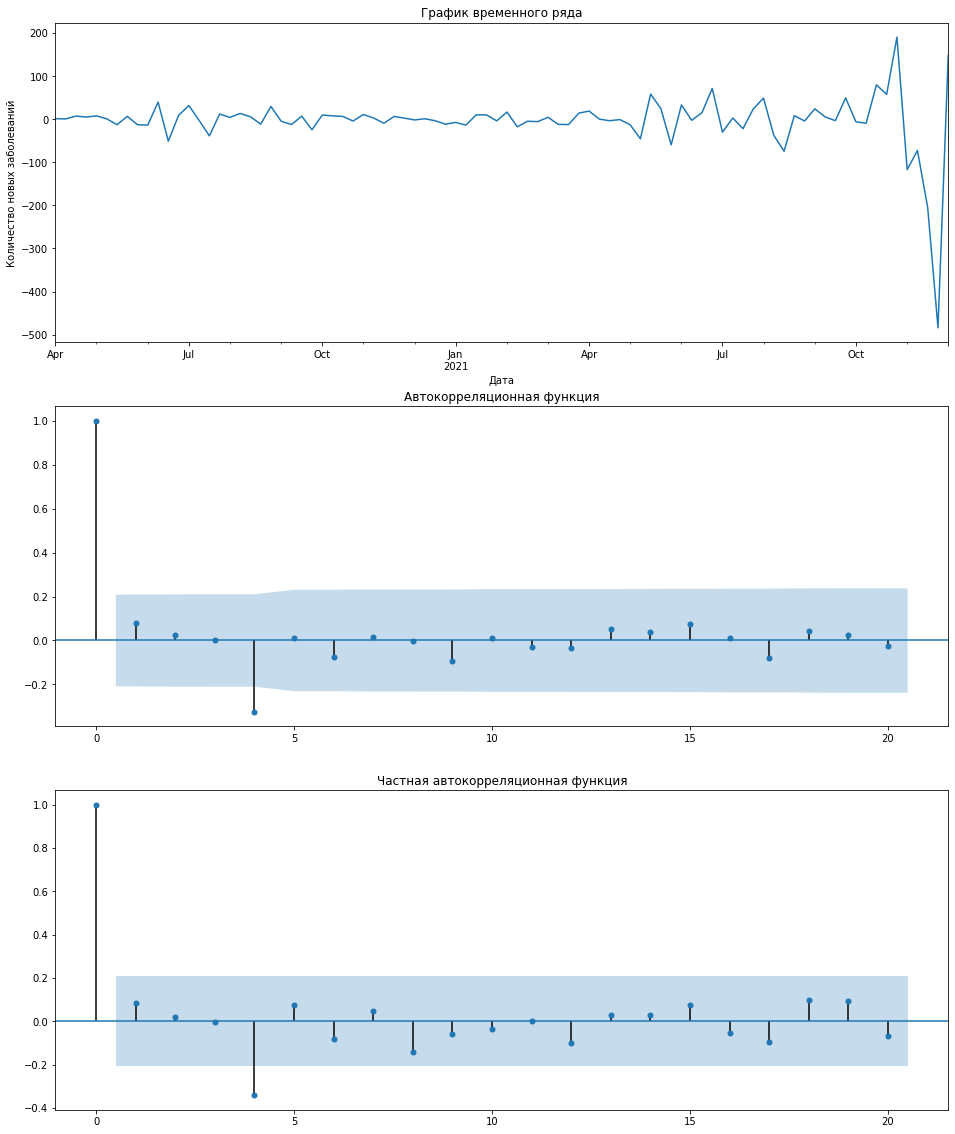


Рисунок 6 – График вторых разностей, их АКФ и ЧАКФ.

Переберем модели с различными значениями параметров p, d, q и выведем 5 с наименьшим значением информационного критерия Байеса (Рисунок 7).

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 7 – результат перебора параметров.

В результате перебора лучшими параметрами оказались p=7, d=1, q=0. Оценим лучшую модель (Рисунок 8).

Изображение выглядит как стол

Автоматически созданное описаниеИзображение выглядит как стол

Автоматически созданное описание

Рисунок 8– оценка наилучшей модели.

Оценка модели АРПСС(7,1,0) выглядит следующим образом :

Построим график предсказанных и исходных значений (Рисунок 9) и рассчитаем функционалы качества модели (Рисунок 10).

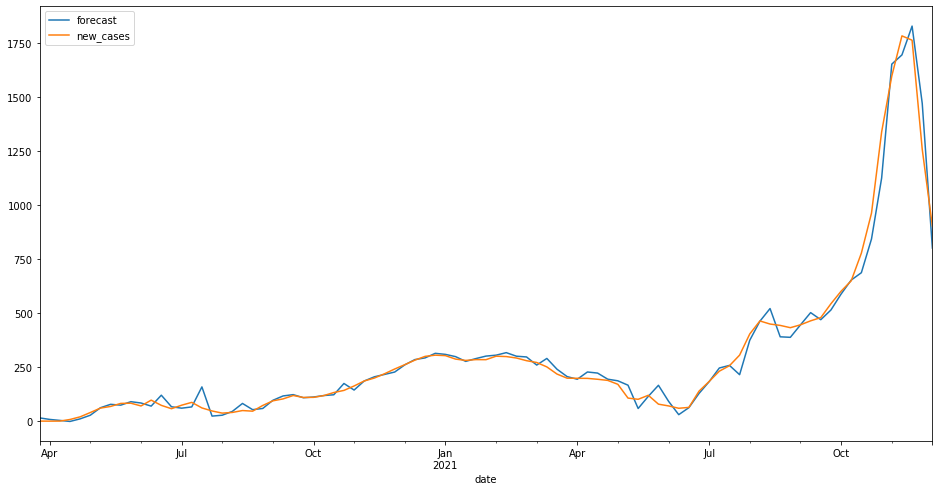


Рисунок 9 - график предсказанных и исходных значений

Изображение выглядит как текст

Автоматически созданное описание

Рисунок 10 – функционалы качества модели

Проверим модель на адекватность. Для этого проверим нормальность распределения остатков модели: проверим их с помощью критерия Дики-Фуллера (рисунок 11), построим их график, АКФ и ЧАКФ (рисунок 12), сравним график распределения остатков с нормальным (рисунок 13) и построим qq-график остатков (рисунок 14)

Изображение выглядит как текст

Автоматически созданное описание

Рисунок 11 – критерий Дики-Фуллера для остатков

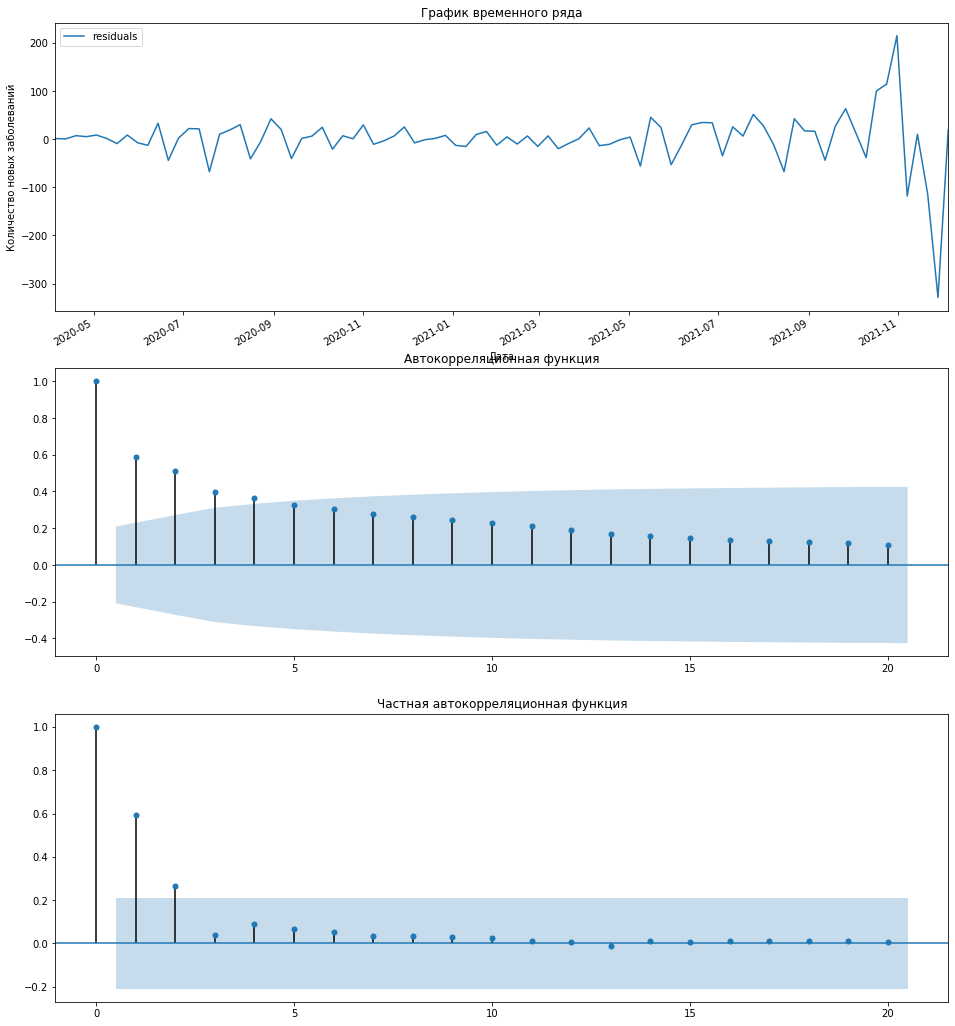


Рисунок 12 – График остатков, их АКФ и ЧАКФ

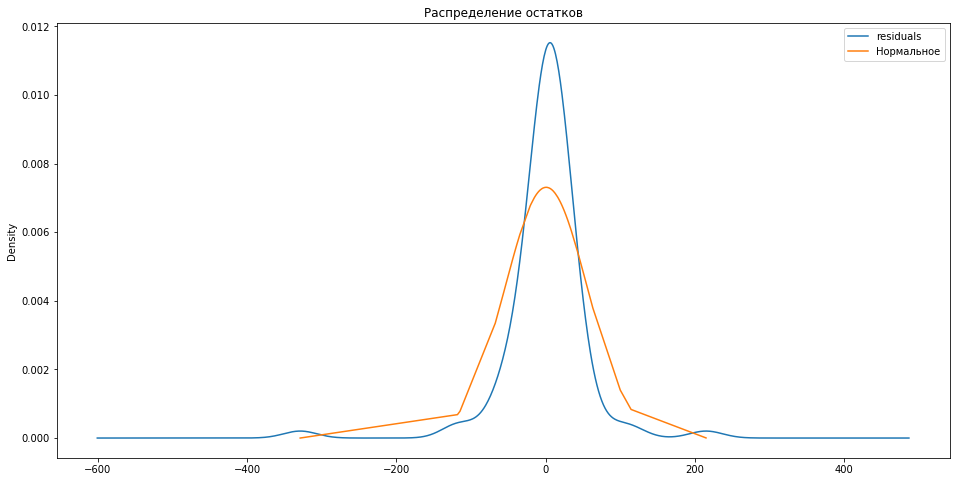


Рисунок 13 – Сравнение распределения остатков с нормальным

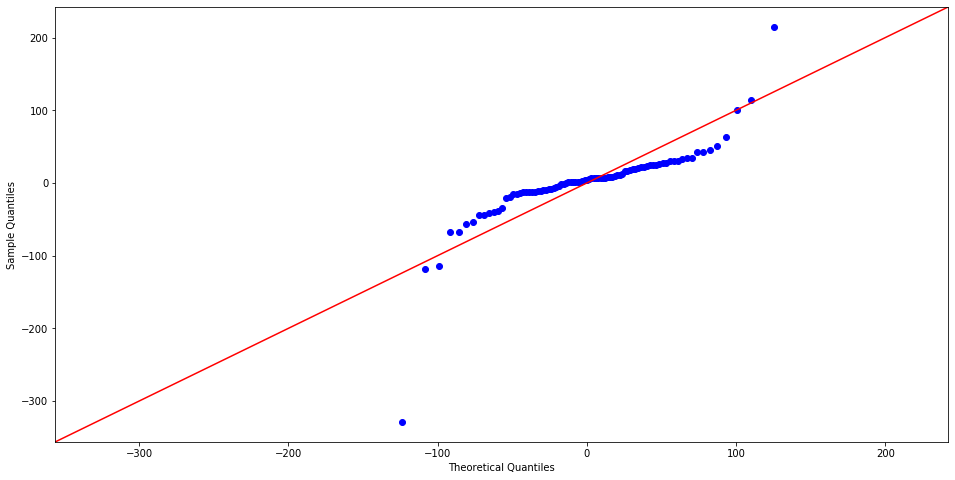


Рисунок 14 – qq-график остатков

Исходя из графиков АКФ и ЧАКФ требование к некоррелированности остатков не выполнено, кроме того, остатки распределены не нормально, следовательно модель не является адекватной.

Спрогнозируем будущие значения ряда на 2 недели вперед (Рисунок 15).

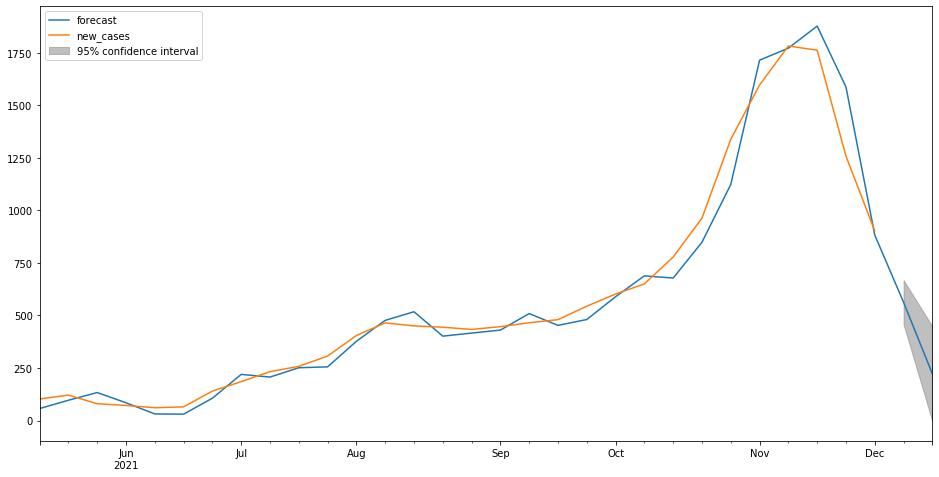


Рисунок 15 – прогноз на 2 недели

Добавим в модель сезонность и построим модель SARIMA. Возьмем сдвиг по сезонности равный 6 месяцам. Проверим модель с помощью критерия Дики-Фуллера (Рисунок 16).

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 16 – Критерий Дики-Фуллера для сезонной модели

Также были построены графики для сезонной модели (Рисунок 17), АКФ и ЧАКФ (Рисунок 18) для D=1.

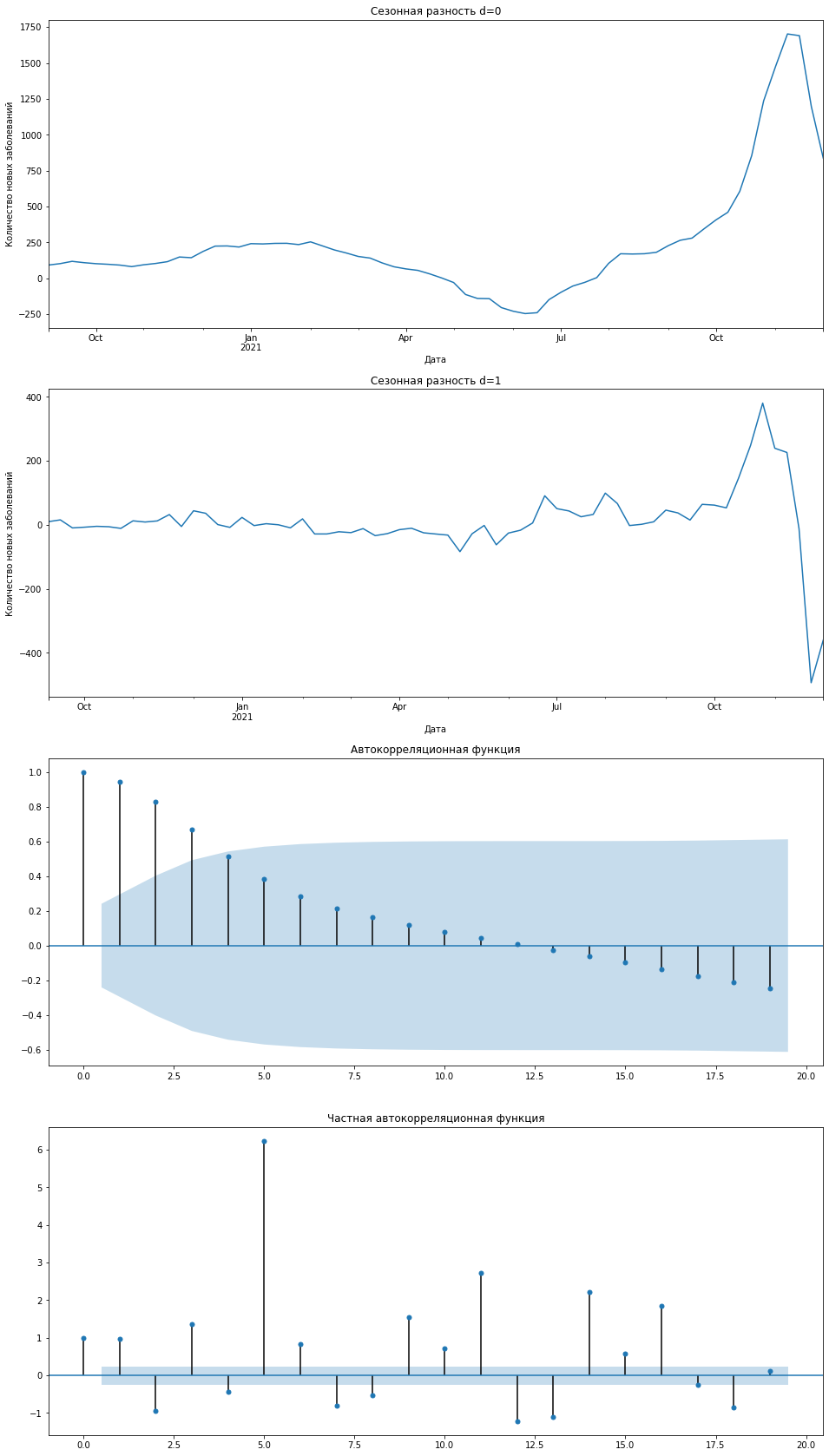


Рисунок 17 – Графики для сезонной модели

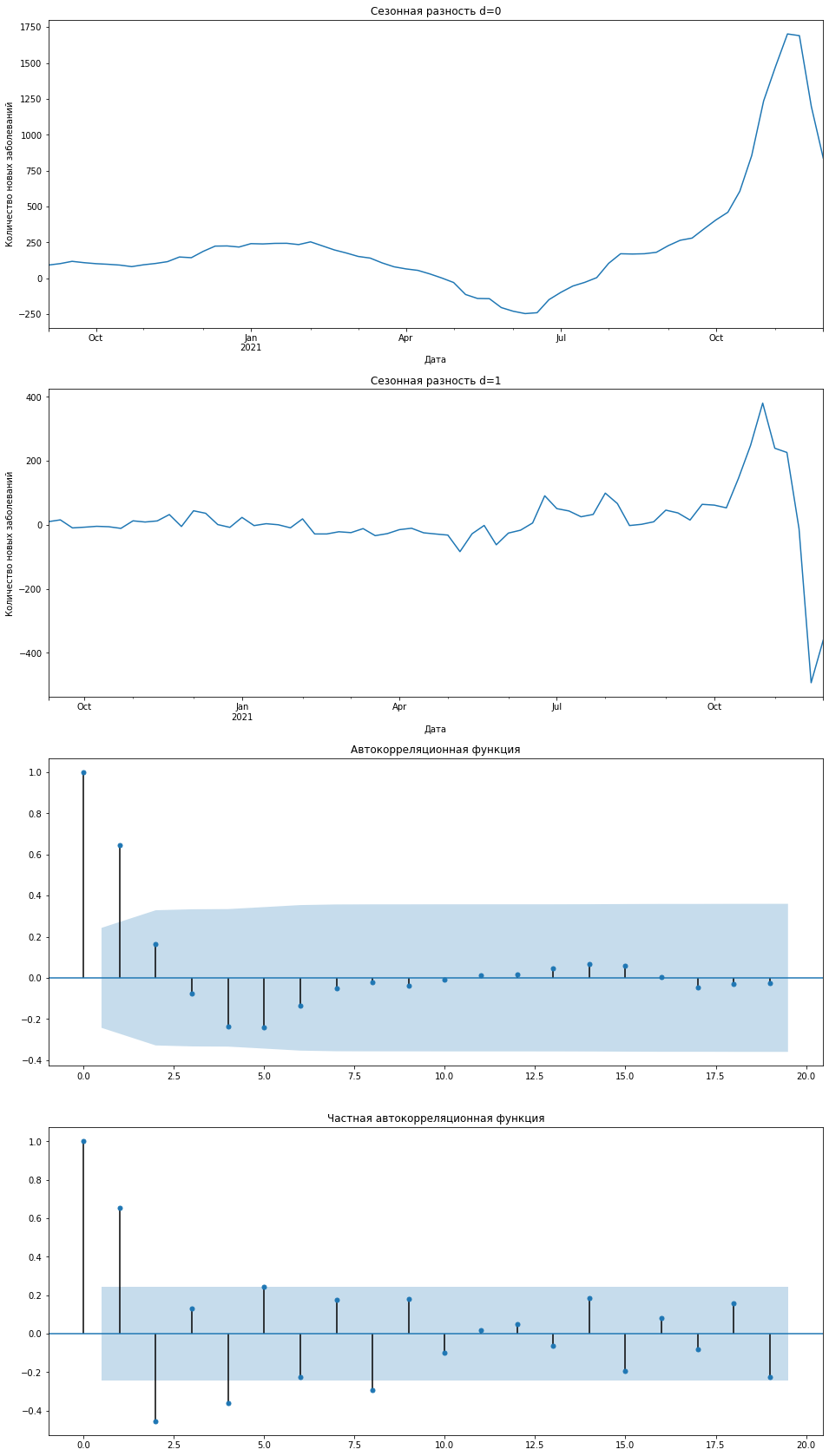


Рисунок 18 – АКФ и ЧАКФ сезонной модели

По графикам АКФ и ЧАКФ определяем P и Q. Выберем P = 2, Q = 1. Используя лучшие значения p, d, q из предыдущей модели строим модель SARIMA (Рисунок 19).

Изображение выглядит как текст, квитанция, снимок экрана

Автоматически созданное описание

Рисунок 19 – оценка модели SARIMA

Проведем анализ остатков. Основная гипотеза Льюнга-Бокса (об отсутствии автокорреляции) принимается. В тесте на гетероскедастичность гипотеза о гомоскедастичности остатков отвергается. Согласно тесту Харке-Бера – остатки не распределены нормально. Построим график значений сезонной модели (Рисунок 20) и рассчитаем функционалы качества (Рисунок 21).

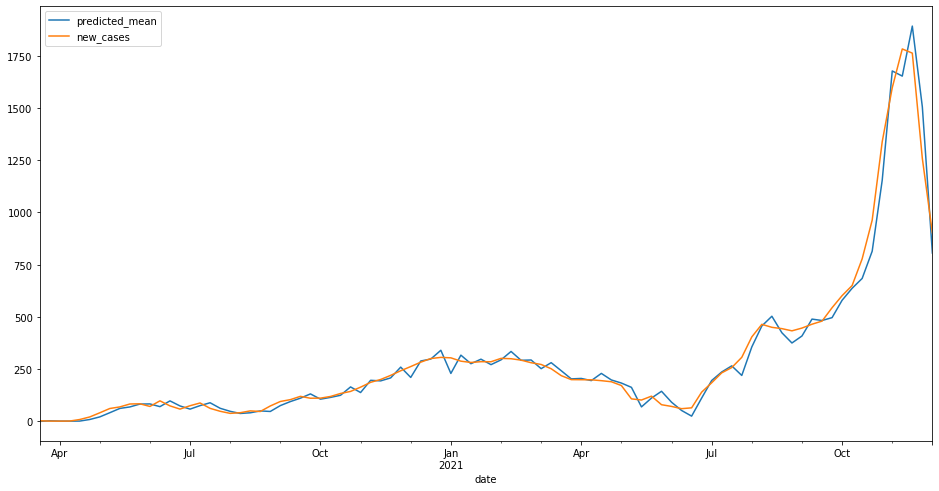


Рисунок 20 – график значений модели SARIMA

Рисунок 21 – функционалы качества модели SARIMA

Изображение выглядит как текст

Автоматически созданное описание

Проверим остатки модели на нормальность. Для этого проверим их с помощью критерия Дики-Фуллера (рисунок 22), построим их график, АКФ и ЧАКФ (рисунок 23), сравним график распределения остатков с нормальным (рисунок 24) и построим qq-график остатков (рисунок 25)

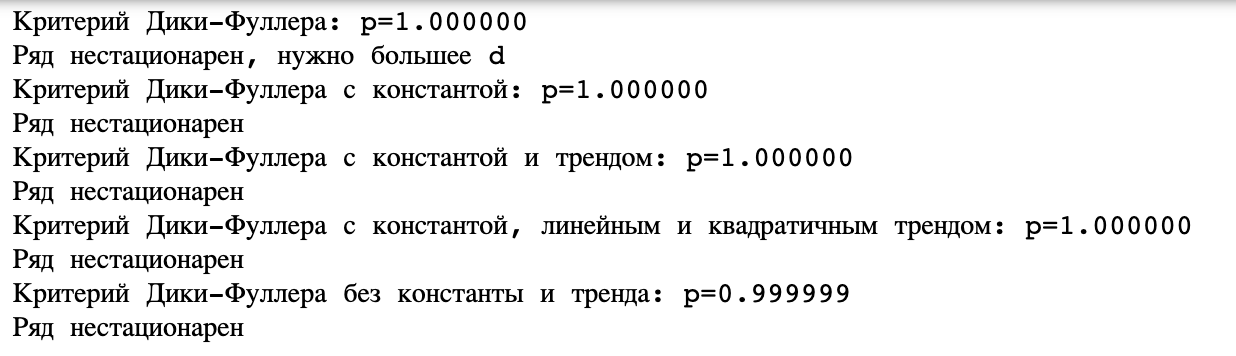


Рисунок 22 – критерий Дики-Фуллера для остатков

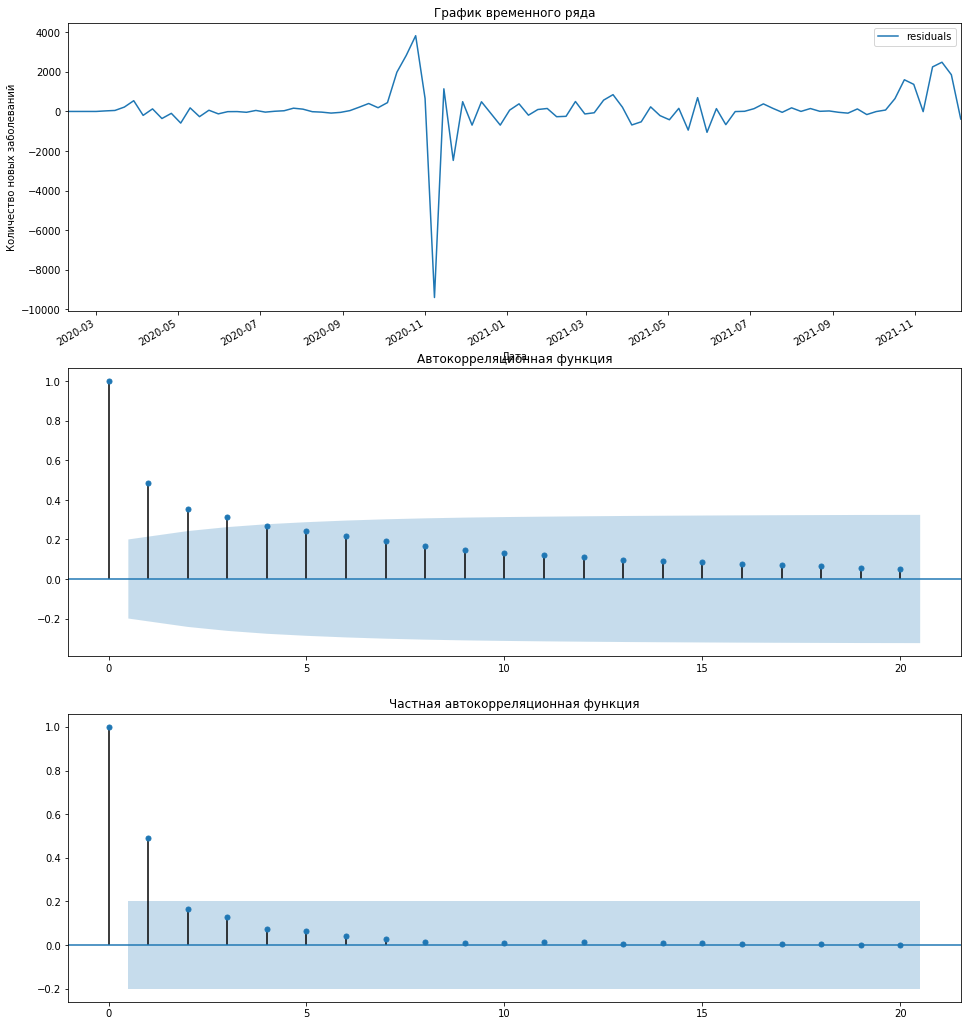


Рисунок 23 – График остатков, их АКФ и ЧАКФ

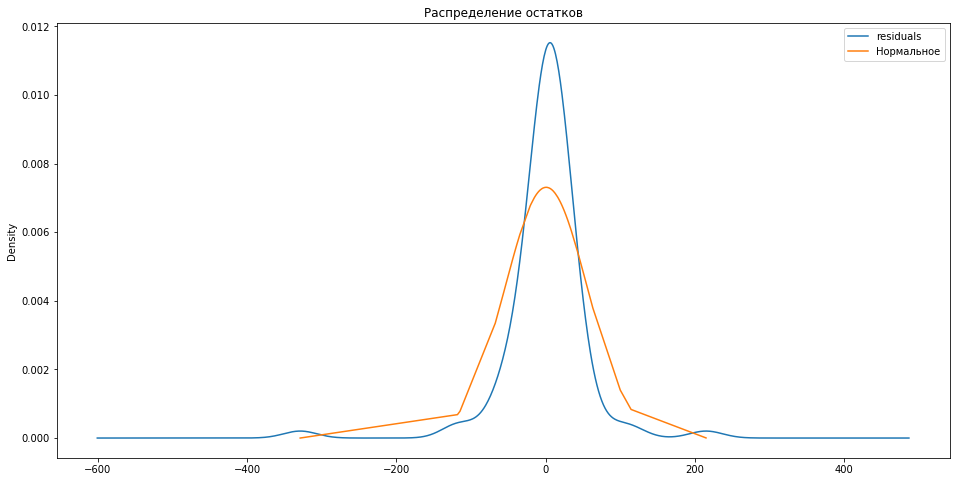


Рисунок 24 – Сравнение распределения остатков с нормальным

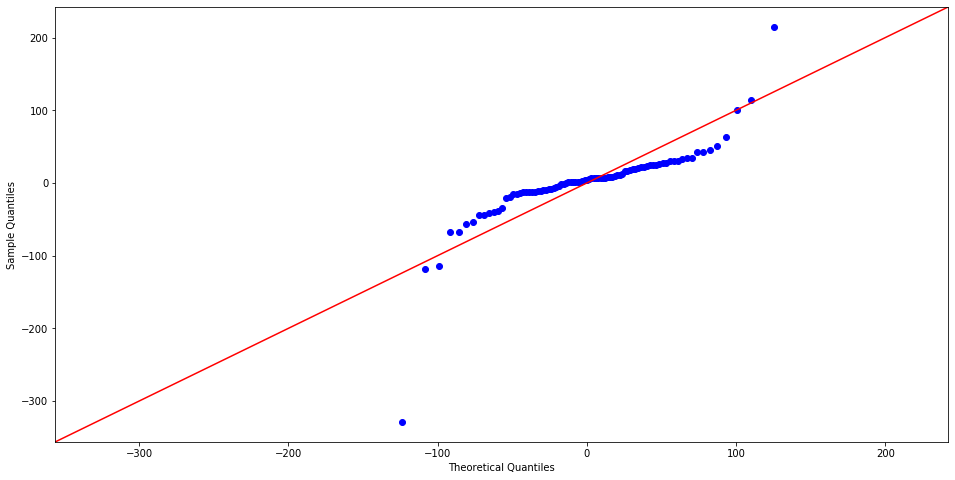


Рисунок 25 – qq-график остатков

Исходя из графиков остатки распределены не нормально, но близко к нему, следовательно модель можно считать адекватной.

Спрогнозируем будущие значения ряда на 2 недели вперед (Рисунок 26).

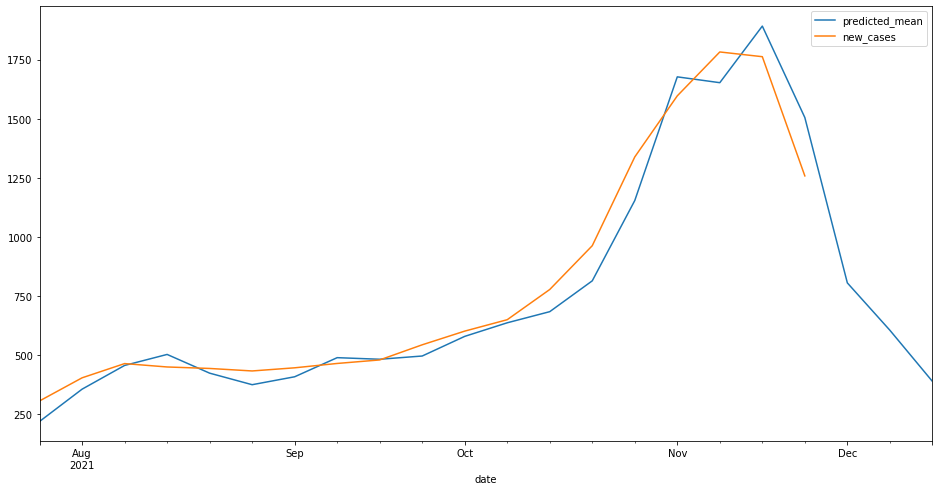


Рисунок 26 – прогноз на 2 недели

Рисунок 27 – график заболеваемости и смертности