Министерство образования и науки Российской Федерации

Федеральное государственное автономное образовательное учреждение   
высшего образования

«Южно-Уральский государственный университет

(Национальный исследовательский университет)»

Институт естественных и точных наук

Кафедра «Математическое и компьютерное моделирование»

Направление 02.03.01 «Математика и компьютерные науки»

КУРСОВАЯ РАБОТА

по дисциплине «Методы прогнозирования и анализа рынка»

Тема: «Прогнозирование динамики финансового состояния кредитной организации ПАО СКБ–БАНК»

Руководитель, доцент каф. МиКМ

\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Акимова А. А.

«\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_2018 г.

Автор работы

Студент группы ЕТ-411

\_\_\_\_\_\_\_\_\_\_\_\_\_Русанова Т. А.

«\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_2018 г.

Работа защищена

с оценкой (прописью, цифрой)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

«\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_2018 г.

Челябинск, 2018

Оглавление

[ВВЕДЕНИЕ 3](#_Toc533367967)

[1. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ВРЕМЕННОГО РЯДА 5](#_Toc533367968)

[**1.1.** **Выявление и устранение аномальных наблюдений** 5](#_Toc533367969)

[**1.2.** **Сглаживание временного ряда** 7](#_Toc533367970)

[**1.3.** **Предварительные расчеты** 9](#_Toc533367971)

[2. АДДИТИВНАЯ МОДЕЛЬ ВРЕМЕННОГО РЯДА 11](#_Toc533367972)

[**2.1.** **Выявление наличия неслучайной составляющей** 12](#_Toc533367973)

[**2.2.** **Построение моделей трендовой составляющей** 15](#_Toc533367974)

[**2.3.** **Выбор адекватной модели трендовой составляющей** 17](#_Toc533367975)

[**2.4.** **Выделение сезонной составляющей** 18](#_Toc533367976)

[***2.5.*** **Построение аддитивной модели** 19](#_Toc533367977)

[3. ЛИНЕЙНАЯ МОДЕЛЬ ВРЕМЕННОГО РЯДА 20](#_Toc533367980)

[**3.1.** **Выявление наличия свойства стационарности временного ряда** 20](#_Toc533367984)

[**3.1.1 Математическое описание** 21](#_Toc533367985)

[**3.1.2 Реализация в пакете EViews** 23](#_Toc533367986)

[***3.2.*** ***Построение авторегрессионной модели AR(p)*** 27](#_Toc533367988)

[***3.2.1.*** ***Построение модели скользящего среднего MA(q)*** 29](#_Toc533367989)

[***3.2.2.*** ***Построение смешанной модели ARMA(p,q)*** 30](#_Toc533367992)

[3.2.3. ***Сравнительный анализ построенных моделей*** 30](#_Toc533367993)

[4. КОМБИНИРОВАННАЯ МОДЕЛЬ ВРЕМЕННОГО РЯДА 33](#_Toc533367994)

[**4.1.** **Построение комбинированной модели** 33](#_Toc533367996)

[5. ТОЧЕЧНЫЙ ПРОГНОЗ ПО ПОСТРОЕННЫМ МОДЕЛЯМ 35](#_Toc533368001)

[**5.1. Точечный прогноз по аддитивной модели** 36](#_Toc533368002)

[**5.2.** **Точечный прогноз по линейной модели** 36](#_Toc533368003)

[**5.3. Точечный прогноз по комбинированной модели** 37](#_Toc533368004)

[**5.4. Сравнительный анализ построенных прогнозов** 37](#_Toc533368005)

[ЗАКЛЮЧЕНИЕ 37](#_Toc533368007)

[СПИСОК ЛИТЕРАТУРЫ 38](#_Toc533368008)

[ПРИЛОЖЕНИЕ А 39](#_Toc533368009)

# ВВЕДЕНИЕ

ПАО «СКБ–БАНК» – крупный сетевой универсальный банк федерального масштаба. Данная кредитная организация входит в топ – 40 банков России.

Акционерный капитал банка составляет 1 822 775 000 руб. и разделён на 1 822 775 000 акций номиналом в 1 рубль. В настоящее время у банка более 35 000 акционеров.

СКБ-банк — один из самых быстрорастущих [банков России](https://ru.wikipedia.org/wiki/%D0%91%D0%B0%D0%BD%D0%BA%D0%B8_%D0%B2_%D0%A0%D0%BE%D1%81%D1%81%D0%B8%D0%B8). К декабрю 2014 г. объём нетто-активов СКБ-банка составил 121 млрд рублей, а величина капитала Банка превысила отметку 14,6 млрд руб.

Портфель кредитов населению на 1 декабря 2014 превысил 65,6 млрд рублей. [Портфель](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D1%80%D1%82%D1%84%D0%B5%D0%BB%D1%8C_(%D1%84%D0%B8%D0%BD%D0%B0%D0%BD%D1%81%D1%8B)) кредитов предприятиям и организациям на 1 декабря 2014 года составил 17,3 млрд рублей. К декабрю 2014 года объём вкладов населения превысил 65,5 млрд рублей.

В ходе курсовой работы исследована динамика финансового состояния ПАО «СКБ–БАНК» средствами технического анализа.

Технический анализ – это прогнозирование изменений значений экономических показателей в будущем на основе анализа изменений этих значений в прошлом. В основе его лежит анализ временных рядов значений и их графиков. Наиболее часто методы технического анализа используются для анализа цен, изменяющихся свободно, например, на биржах.

Случайный процесс (СП) – параметризированное семейство случайных величин {Y()}, где параметр – время.

СП называется СП с непрерывным временем, если система, в которой он протекает, меняет свои состояния в любой момент времени.

Временной ряд – набор упорядоченных во времени случайных величин:

{Y()} = Y(), Y(), …, Y(),

т.е. наблюдений над случайным процессом с непрерывным временем Y() в моменты времени , где – моменты времени и i=1, 2, …, n.

– наблюдение над случайной величиной Y(), т.е. конкретная реализация не является случайной величиной. Далее наблюдения будем называть фактическими наблюдениями. Фактические значения анализируемого временного ряда приведены в Приложении А.

Методы, описанные в курсовой работе, предназначены для выявления структуры временных рядов и для их прогнозирования. Прогноз будущих значений временного ряда используется для эффективного принятия решений.

Актуальность работы заключается в большой роли финансовых организации в современной экономике, интересе к ним больших групп людей и в занимаемом в мировой экономике месте ПАО «СКБ–БАНК».

# ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ВРЕМЕННОГО РЯДА

На этапе предварительной обработки производится построение и изучение графика временного ряда (диаграммы рассеяния). Графические методы анализа позволяют сделать предварительные выводы о характере процесса, которые затем могут быть проверены и уточнены с помощью расчета конкретных характеристик ряда.

В разделе 1.1 временной ряд будет проверен на наличие аномальных наблюдений с помощью критерия Ирвина. В разделе 1.2 мы рассмотрим экспоненциальное сглаживание временного ряда. В разделе 1.3 проведем предварительные расчеты: найдем числовые характеристики временного ряда, построим коррелограмму.

## **Выявление и устранение аномальных наблюдений**

Под аномальным наблюдением понимается отдельное значение временного ряда, которое не отвечает потенциальным возможностям исследуемой экономической системы и которое, оставаясь в качестве уровня ряда, оказывает существенное влияние на значение основных характеристик временного ряда.

Причинами аномальных явлений могут быть ошибки технического порядка, а также воздействия факторов, имеющих объективный характер, но проявляющихся эпизодически. Аномальные значения подлежат выявлению и устранению.

**Критерий Ирвина**

Критерий Ирвина применяют для выявления аномальных наблюдений (грубых ошибок) выборки из нормально распределённой случайной величины.

Для всех (или только “подозреваемых” в аномальности) наблюдений можно сформулировать следующие статистические гипотезы:

i –е наблюдение не является аномальным

i –е наблюдение является аномальным

Для проверки гипотез найти значение критерия:

Так же необходимо вычислить предельное значение критерия . Предельное значение критерия зависит от числа наблюдений n и вероятности𝛼, Таблица 1.1.

Таблица 1.1 – Предельные значения

|  |  |  |
| --- | --- | --- |
| N | 𝛼=0,05 | 𝛼=0,01 |
| 2 | 2,8 | 3,7 |
| 3 | 2,2 | 2,9 |
| 10 | 1,5 | 2,0 |
| 20 | 1,3 | 1,8 |
| 30 | 1,2 | 1,7 |
| 50 | 1,1 | 1,6 |
| 100 | 1,0 | 1,5 |
| 400 | 0,9 | 1,3 |

Если n находится между табличными значениями и , то необходимо выполнить линейную интерполяцию:

Если , то с вероятностью 𝛼 ошибки первого рода принимается гипотеза .

Для исходных данных значения статистик Ирвина лежат в диапазоне от 0,04 до 1,56. Согласно критерию Ирвина, с вероятностью 5% во временном ряду присутствуют аномальные уровни, так как не все , где .

## **Сглаживание временного ряда**

Цель сглаживания временного ряда заключается в получении ряда с меньшим разбросом уровней, что в ряде случаев позволяет на основе визуального анализа сделать вывод о наличии тенденции и ее характерных особенностях.

Весьма эффективным и надежным методом является метод экспоненциального сглаживания. Основные достоинства метода состоят в возможности учета весов исходной информации, в простоте вычислительных операций, в гибкости описания различных динамик процессов. Метод экспоненциального сглаживания дает возможность получить оценку параметров тренда, характеризующих не средний уровень процесса, а тенденцию, сложившуюся к моменту последнего наблюдения. Наибольшее применение метод нашел для реализации среднесрочных прогнозов. Для метода экспоненциального сглаживания основным и наиболее трудным моментом является выбор параметра сглаживания α и начальных условий.

Метод экспоненциального сглаживания выполняется по рекуррентной формуле (модель Брауна):

где элемент ряда, параметр сглаживания, ,

— значение экспоненциально взвешенной средней в момент ,

– начальная экспоненциально взвешенная средняя.

Важную роль в методе экспоненциального сглаживания играет выбор оптимального параметра сглаживания α , так как именно он определяет оценки коэффициентов модели, и выбора параметра .

В зависимости от величины параметра α прогнозные оценки по-разному учитывают влияние исходного ряда наблюдений: чем больше α, тем больше вклад последних наблюдений в формирование тренда, а влияние начальных условий быстро убывает. При малом α прогнозные оценки учитывают все наблюдения, при этом уменьшение влияния более «старой» информации происходит медленно. Известны два основных соотношения, позволяющие найти приближенную оценку α. Первое соотношение Брауна:

где n – число точек ряда. Вторым является соотношение Мейера:

где — среднеквадратическая ошибка модели, - среднеквадратическая ошибка исходного ряда. Однако использование последнего соотношения затруднено тем, что достоверно определить из исходной информации очень сложно.

Для выбора начального значения также используется два метода. Если есть данные о развитии явления в прошлом, то ­— это среднее арифметическое значение из всех имеющихся уровней временного ряда или из какой-то их части. В случае если таких значений нет, присваивается значение первого элемента временного ряда.

Для применения экспоненциального сглаживания к временному ряду примем (по соотношению Брауна при). Для примем значение . Воспользуемся средствами MS Excel, используя надстройку «Анализ данных» и средство анализа «Экспоненциальное сглаживание». Тогда ряд примет вид (см. Таблицу 1.2)

Таблица 1.2 – Изменение временного ряда

|  |  |  |
| --- | --- | --- |
| Дата | Значения после применения критерия Ирвина | Значения после применения экспоненциального сглаживания |
| 2 кв. 2018 | 416285 | 416285 |
| 1 кв. 2018 | 385878 | 416285 |
| 4 кв. 2017 | 396161 | 412332 |
| 3 кв. 2017 | 406444 | 410230 |
| 2 кв. 2017 | 416727 | 409738 |
| 1 кв. 2017 | 92097 | 410646 |
| 4 кв. 2016 | -716318 | 369235 |
| 3 кв. 2016 | -1174833 | 228113 |
| 2 кв. 2016 | 406003 | 45730 |
| 1 кв. 2016 | 119011 | 92566 |
| 4 кв. 2015 | 1017501 | 96003 |
| 3 кв. 2015 | 845311 | 215798 |
| 2 кв. 2015 | 211148 | 297635 |
| 1 кв. 2015 | 345155 | 286392 |

Рассмотренный метод является одним из наиболее надежных и широко применяется в практике прогнозирования. Одно из наиболее перспективных направлений развития данного метода представляет собой метод разностного прогнозирования, в котором само экспоненциальное сглаживание рассматривается как частный случай.

## **Предварительные расчеты**

**Числовые характеристики временного ряда**

В случае аддитивной модели:

Математическое ожидание оценивается как **выборочное среднее**:

Дисперсия оценивается как **выборочная дисперсия**:

Степень статистической связи между последовательностями и , сдвинутыми относительно друг друга на *l* моментов времени, т.е. с лагом (порядком коэффициента автокорреляции) *l*, может быть определена с помощью коэффициента автокорреляции

Коэффициент автокорреляции *p(l)* оценивается как **выборочный коэффициент автокорреляции:**

Нормированная автокорреляционная функция временного ряда – последовательность коэффициентов автокорреляции *r(0), r(1), r(2),….*

Для анализа значений нормированной автокорреляционной функции удобно использовать график, который называется *коррелограммой*. Она изображает зависимость значений коэффициентов автокорреляции от величины лага.

Частный коэффициент автокорреляции измеряет связь между текущим значением ряда и его предыдущими значениями , когда влияние всех промежуточных лагов устранено.

Математическое ожидание и дисперсию найдем с помощью Excel, а все остальные числовые характеристики временного ряда – в пакете E’Views.

Расчетные значения для характеристик представлены в таблице 1.3.

Таблица 1.3 – Расчетные значения

|  |  |
| --- | --- |
| Характеристика | Значение |
| Математическое ожидание |  |
| Дисперсия |  |

На Рисунке 1.1 слева показана коррелограмма, справа – частная коррелограмма.

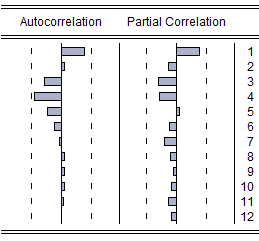


Рисунок 1.1 – коррелограмма(слева) и частная коррелограмма(справа)

Значения, полученные в ходе предварительной обработки, используются в дальнейших расчетах.

# АДДИТИВНАЯ МОДЕЛЬ ВРЕМЕННОГО РЯДА

Для анализа структуры временных рядов используют аддитивную модель, которая имеет следующий вид:

Y(τi) = q(τi) + Ԑ(τi),

здесь Ԑ(τi) – случайная составляющая, которая отражает случайные колебания или шумы процесса,

q(τi) – неслучайная (систематическая) составляющая, которая может включать одну или несколько из следующих компонент:

t(τi) – трендовая составляющая, наблюдаемая в течение длительного периода времени, которая отражает влияние долговременных факторов,

s(τi) – сезонная компонента, которая описывает регулярные периодические колебания,

p(τi) – периодическая (циклическая) компонента, которая описывает длительные периоды (более одного года) относительного подъема и спада.

Мы будем строить аддитивную модель, включающую только трендовую и сезонную составляющие.

Прежде всего, в разделе 2.1 мы покажем, что временной ряд имеет неслучайную составляющую. Далее, в разделе 2.2, построим 2 модели трендовой составляющей, после чего в разделе 2.3 сделаем выбор лучшей модели трендовой составляющей, и в разделе 2.4 выделим сезонную составляющую. В последнем разделе 2.5, на основании полученных результатов, построим аддитивную модель.

## **Выявление наличия неслучайной составляющей**

В аддитивной модели отсутствует неслучайная составляющая M() = const. (т.е временной ряд состоит из статистических независимых наблюдений, случайно варьирующихся около некоторого постоянного значения).

Таким образом, вместо проверки наличия неслучайной составляющей будем проверять постоянство математического ожидания. Сформулируем две статистические гипотезы:

Рассмотрим два критерия для проверки этих гипотез.

**Критерий 1 (Метод доверительных разностей средних уровней)**

Для проверки неслучайной составляющей нужно сделать следующее:

1. Разбить временной ряд на две части и , где

*i = 1,2,…,;*

*j = +1, +2, …, +;*

1. Предположить, что и имеют нормальное распределение.
2. Найти для обеих частей выборочные средние и выборочные дисперсии.

После вычислений получим, что и .

Проверить равенство дисперсий (по Критерию Фишера). Если выполнено неравенство

то дисперсии одинаковы с уровнем значимости , поэтому можно применять критерий Стьюдента.

Если неравенство не выполнено, дисперсии неодинаковы, нужно либо применить другой критерий (не Критерий 1), либо принять гипотезу .

неравенство выполняется, поэтому дисперсии равны и можем применить Критерий Стьюдента.

1. Критерий Стьюдента (проверить равенство математических ожиданий). Если выполнено

где

в Excel , то гипотеза принимается с уровнем значимости .

Идея критерия: если выборочные средние , то тренда во временном ряде нет, иначе – тренд есть.

В нашем случае , следовательно принимается гипотеза . Таким образом, согласно Критерию 1, неслучайная составляющая во временном ряде отсутствует.

**Критерий 2 (Критерий Серий)**

Критерий серий необходим для проверки постоянства математических ожиданий. Таким образом, он позволяет проверить, является ли случайным порядок появления двух значений переменной. Серия – это последовательность похожих наблюдений. Если в выборке либо слишком много серий, либо слишком мало, то эта выборка не является случайной. Для проверки выборки на случайность необходимо сделать следующее.

1. Упорядочить члены временного ряда по возрастанию:

.

1. Найти выборочную медиану:
2. По элементам , расположенным согласно исходному временному ряду и отличным от , образовать последовательность серий по правилу:

.

Полученная нами последовательность плюсов и минусов характеризуется общим числом серий  и протяженностью самой длинной серии . Под «серией» понимается последовательность подряд идущих плюсов или подряд идущих минусов (в частном случае серия может состоять только из одного плюса или только из одного минуса и тогда ее протяженность равна единице). Очевидно, что если наблюдения стохастически независимы (выборка случайна), то чередование плюсов и минусов в последовательности должно быть более или менее «случайным», т. е. эта последовательность не должна содержать слишком длинных серий подряд идущих плюсов или подряд идущих минусов, и соответственно общее число серий  не должно быть слишком малым.

**Вывод критерия серий:** если хотя бы одно из неравенств

не выполнено, то гипотеза отвергается с вероятностью ошибки , где . Здесь – взятие целой части числа z.

Второе неравенство не выполняется, отсюда следует, что чередование «+» и «–» неслучайно, значит, отвергается гипотеза , принимается гипотеза .

Таким образом, согласно Критерию 2, неслучайная составляющая имеет место. Следовательно, в дальнейшем имеет смысл выделения неслучайной составляющей.

## **Построение моделей трендовой составляющей**

Тренд – это аналитическая функция, которая описывает тенденцию изменения явления и связывает единым законом развития все последующие уровни ряда динамики.

Трендовая составляющая t (τ) отражает устойчивую и долговременную тенденцию временного ряда.

Таким образом, возникает задача выделения тренда, т.е. построения оценки для функции t(τ) (или оценок (τi) для значений t (τi)) по заданной временной выборке {τi, yi}. При этом предполагается, что остальные составляющие p(τ), s(τ) временного ряда отсутствуют.

Полагая время τ независимой переменной, оценим функцию t (τ), используя метод парной регрессии, который основан на аддитивной модели временного ряда:

Y(τi) = t(τi) + ε(τi),

где случайные величины ε(τi) удовлетворяют условиям Гаусса-Маркова:

1) дисперсии δi – конечные, одинаковые, независимые от измерений и таковы, что

M(ε(τi)) = 0,

M(ε(τi), ε(τj)) = ,

2) случайные величины ε(τi) имеют нормальное распределение N (0, δ2).

Для того, чтобы определить характер тренда, необходимо выбрать вид функции t(τ). Предварительно мы анализируем графическое изображение ряда, т.е. строим диаграмму рассеивания по точкам {τi, yi}. С помощью диаграммы рассеивания можно сделать выбор вида функции из представленных ниже.

1. Линейная функция:

t(τ) = β0 + β1τ

используется для представления процессов с постоянной скоростью изменения.

1. Полиномиальная

После выбора вида функции t(τi) построим уравнение регрессии (τ), зависящее от коэффициентов b0, b1,…,bk, которые являются оценками коэффициентов β0, β1,…, βk функции тренда t (τ).

Вычислим коэффициенты b0, b1,…,bk по методу наименьших квадратов, минимизируя отклонение расчетных значений от фактических:

F(b0, b1,…,bk) = .

Для этого нужно найти частные производные 1 порядка по каждому из коэффициентов bi, i=, приравнять к нулю и решить полученную систему из k уравнений.

Получив коэффициенты b0, b1,…,bk, примем составленное уравнение регрессии (τ) в качестве оценки для функции тренда t(τ), которая может быть использована для дальнейшего анализа временного ряда или его прогнозирования.

В результате визуального анализа диаграммы рассеивания, представленной на рисунке 2.1, для определения характера тренда выбраны линейная и полиномиальная (4–ей степени) функции.

Рисунок 2.1 – Диаграмма рассеивания

С помощью встроенных функций Exсel найдем коэффициенты уравнений регрессии. Таким образом, мы получаем следующие модели трендовой составляющей:

1. линейная функция:

t(τ) = –10827τ + 382745,

1. полиномиальная функция:

## **Выбор адекватной модели трендовой составляющей**

Оценка точности модели тенденции заключается в оценке близости модельных значений тенденции к фактическим уровням ряда и осуществляется с помощью вычисления индекса детерминации.

Введем суммы:

; ,

где

Индекс детерминации показывает, какая доля изменения временного ряда обусловлена изменением переменной , характеризует близость уравнения тренда к исходным данным, которые содержат «нежелательную» случайную составляющую .

В случае линейного тренда справедливо тождество , где – коэффициент детерминации линейной регрессии:

Если , то справедлива гипотеза : нелинейную регрессию можно заменить линейной.

Индекс детерминации можно использовать только тогда, когда значения чисел коэффициентов регрессии одинаково. При различных значениях необходимо использовать приведенный индекс детерминации.

Приведенный индекс детерминации

где m – число коэффициентов в уравнении регрессии.

Необходимо посчитать все и выбрать из них наибольшее, которое соответствует лучшему уравнению регрессии.

Для линейной модели , поэтому в дальнейшем будем использовать полиномиальную модель 4-ой степени, для которой .

## **Выделение сезонной составляющей**

Сезонность–периодическое и связанное с календарным периодом отклонение от тренда. Сезонная составляющая временного ряда описывает регулярные колебания малого периода (меньше года).

Рассмотрим два метода выявления сезонной составляющей.

**Метод 1.** Пусть период сезонности равен τ (для ежеквартальных данных *τ*=4), *h=n/τ* (число наблюдений по каждому периоду). Оценим сезонность как разность между средним по всем одноименным периодам или кварталам и средним по всем данным:

**Метод 2.** Необходимо сгладить временной ряд (m=n/*τ)* по следующей формуле:

Сезонность оценивается как отклонения фактических наблюдений от сглаженных значений по формулам:

В таблице 2.1 представлены значения сезонной составляющей финансового состояния ПАО «СКБ БАНК» рассчитанные методами 1 и 2.

Таблица 2.1 – Выявление сезонной составляющей по методам 1 и 2

|  |  |  |
| --- | --- | --- |
| № | Метод 1 | Метод 2 |
| 1 квартал | -76937,07 | -145394,07 |
| 2 квартал | 83570,18 | 19487,04 |
| 3 квартал | -482313,24 | -187241,00 |
| 4 квартал | -332902,57 | -394035,19 |

Проведённые исследования сезонной составляющей позволяют сделать вывод о том, что прирост процентного дохода СКБ БАНК наблюдается во втором квартале.

* 1. **Построение аддитивной модели**

Таким образом, после проведения всех расчетов временной ряд содержит неслучайную составляющую, общий вид аддитивной модели:

Y() = +,

где – трендовая составляющая, наблюдаемая в течение длительного периода времени, которая отражает влияние долговременных факторов;

– сезонная компонента, которая описывает регулярные периодические колебания.

В разделе 2.2 было найдено уравнение трендовой составляющей, тогда аддитивная модель будет иметь следующий вид:

где значения приведены в таблице 5.1.



# ЛИНЕЙНАЯ МОДЕЛЬ ВРЕМЕННОГО РЯДА

Временной ряд называется стационарным в широком смысле (далее – стационарным), если числовые характеристики случайных величин Y() не зависят от момента времени .

Методы определения наличия свойства стационарности данного ВР приведены в разделе 3.1. К таким методам относят следующие методы:

1. DF (Dickey Fuller);

2. ADF (Advanced Dickey Fuller);

3. KPSS (Kwiatkowki Phillips Schmidt Shin).

Проверка наличия стационарности временного ряда позволяет определить класс моделей, которые могут быть построены для данного временного ряда. А именно, в разделах 3.2 – 3.2.3 – модели AR(p), MA(q), ARMA(p,q) стационарных временных рядов.

3. 1. **Выявление наличия свойства стационарности временного ряда**

### **3.1.1 Математическое описание**

Будем говорить, что временной ряд имеет порядок интегрирования k т.е. , если сам временной ряд и все его ряды разностей порядка до k-1 включительно не стационарны, а ряд разностей порядка k - стационарен.

Здесь разность порядка k определяется как , где разность первого порядка .

Стационарный временной ряд имеет порядок интеграции ноль.

Последовательно проводится проверка первоначального временного ряда, ряда первых разностей, ряда вторых разностей и т.д., что позволяет определить порядок интегрирования временного ряда.

Для того, чтобы проверить наш временной ряд на стационарность, принято использовать несколько тестов. Основная идея тестов заключается в том, чтобы по данным временного ряда подобрать параметры одной из известных моделей, после чего, по значениям параметров модели, сделать вывод о наличии стационарности.

Проверку на стационарность будем проводить в пакете EViews.

Рассмотрим несколько тестов:

1. Тест Дикки-Фуллера(DF);
2. Расширенный Тест Дики-Фуллера(ADF);
3. Тест Квятковского-Филлипса-Шмидта-Шина(KPSS).
4. DF-тест основан на оценке параметра приведенного уравнения , где

- фактические наблюдения временного ряда,

- ошибка.

Выдвинем гипотезы:

- наш временной ряд не стационарен.

- наш временной ряд стационарен.

Если значение t-статистики Стьюдента для параметра меньше табличного(критического) значения DF-статистики, то принимается гипотеза , иначе принимается гипотеза .

Таблицы Дики-Фуллера (значения DF-статистики) рассчитаны для уровней значимости в 1, 5, 10 %.

DF-тест применим только если ряд случайных составляющих не автокоррелирован. Если данное условие нарушено, то необходимо включить разность с необходимым количеством лагов. Именно в этом случае применяется расширенный тест Дики-Фуллера.

2. Расширенный тест Дики-Фуллера (ADF-тест) имеет такие же критические значения.

Модель расширенного теста Дики-Фуллера имеет вид:

.

Например, модель расширенного теста Дики-Фуллера, использующая AR (2), имеет вид:

.

Данную модель можно представить в виде:

.

Для определения количества включаемых лагов необходимо при выборке объемом менее 81 наблюдения - 2 лага, при выборка объемом от 81 до 256 наблюдений - 3 лага, и т.д.

Вывод о наличии стационарности по ADF-тесту производится так же, как и в случае DF-теста.

3. Тест стационарности KPSS рассматривает ряд вида

где - стационарный процесс и - случайное блуждание, определяемое как ,

- нормально распределенная случайная величина с нулевым математическим ожиданием, и - коэфициент.

Выдвигаются две гипотезы:

: временной ряд стационарен ( или ),

: временной ряд не стационарен (.

Вывод о наличии стационарности по тесту KPSS производится так же, как и в случае DF-теста.

### **3.1.2 Реализация в пакете EViews**

Для того, чтобы проверить наличие стационарности временного ряда, воспользуемся статистическим пакетом EViews. Данный пакет позволяет нам быстро и наглядно получить искомые характеристики временного ряда.

Для выявления стационарности необходимо взять наш временной ряд с привязкой к дате, пример изображен на рисунке 3.1.

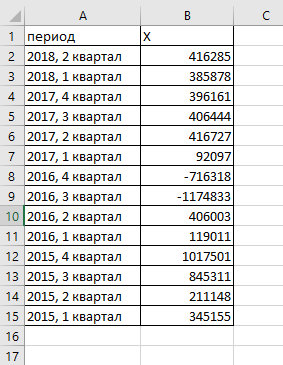


Рисунок 3.1 – Временной ряд с привязкой к дате

В пакете EViews временной ряд с привязкой к дате будет представлен в виде двух окон с таблицами, показанных на рисунке 3.2.

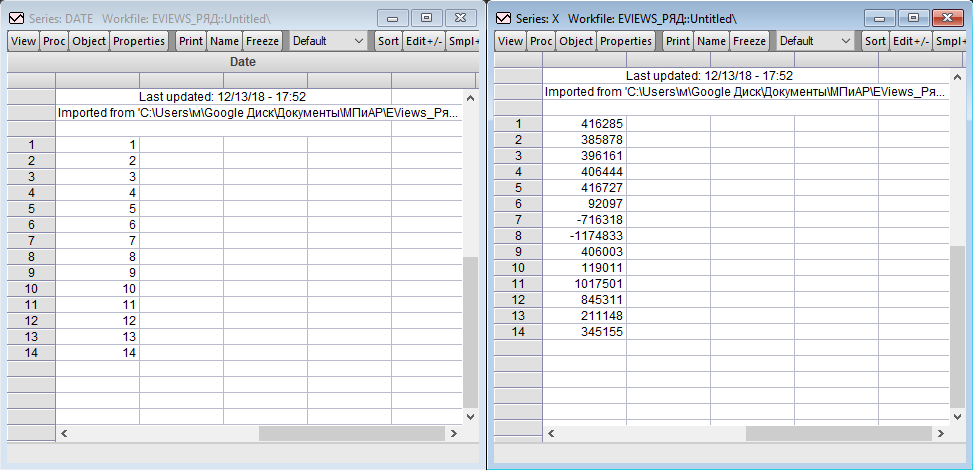


Рисунок 3.2 – Представление временного ряда в пакете EViews

Встроенными функциями пакета EViews представим наш временной ряд графически, порядок действий приведен на рисунке 3.3.

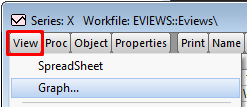


Рисунок 3.3 – Порядок представления временного ряда в графическом виде

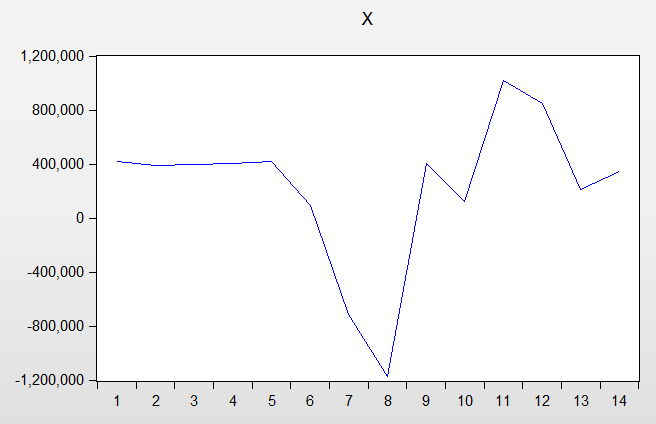


Рисунок 3.4 – Диаграмма рассеяния временного ряда

По полученной диаграмме рассеяния можно сделать вывод о том, что наш временной ряд не стационарен, так как присутствует тренд. Чтобы проверить стационарность используем тесты.

Пакет EViews предлагает несколько тестов на наличие стационарности. Мы рассмотрим следующие: тест Дики-Фуллера, расширенный тест Дики-Фуллера, тест KPSS (Кватковского- Филлипса-Шмидта-Шина).

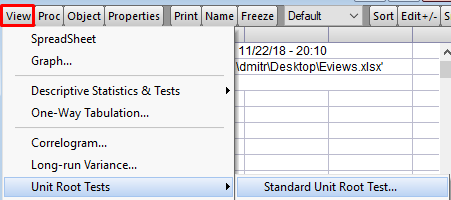


Рисунок 3.5 – Порядок проведения тестов

Для тестирования временного ряда на стационарность с использованием обычного теста Дики-Фуллера в появившемся окне выбрать параметр User specified, равный 1. (см. рисунок 3.6).

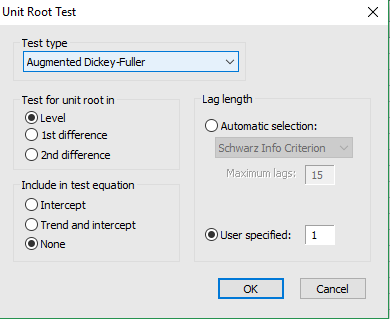


Рисунок 3.6 – Окно выбора параметров для проведения теста Дики-Фуллера

Результаты представлены на рисунке 3.7.



Рисунок 3.7 – Результаты теста Дики-Фуллера

На рисунке 3.7 приведено расчетное значение статистики Дики-Фуллера для стандартного теста -1.806557. Также приведено значение Prob. 0.0687, а также критические значения этой статистики на 1, 5 и 10%-м уровнях значимости -2.771926, -1.974028, -1.602922 соответственно.

В силу того, что Prob., принимающее значение в диапазоне от 0 до 1, находится ближе к нулю, то по этому параметру можно предположить, что временной ряд стационарен.

Проведем расширенный тест Дики-Фуллера для ряда первых разностей, см. рисунок 3.8.

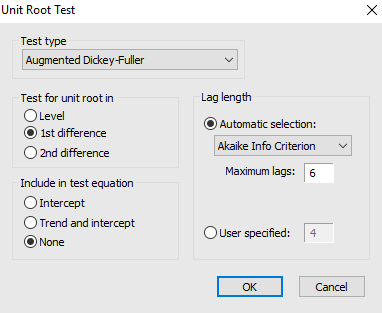


Рисунок 3.8 – Окно выбора параметров для проведения теста Дики-Фуллера по ряду первых разностей



Рисунок 3.9 – Результаты расширенного теста Дики-Фуллера по ряду первых разностей

По результатам теста Prob. равно 0,1925, следовательно, ряд первых разностей стационарен.

Следуя инструкциям, представленным на рисунке 3.10 проведем тест KPSS для временного ряда.

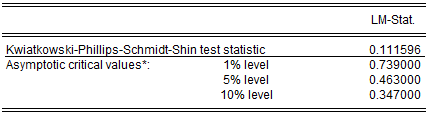


Рисунок 3.10 – Результаты теста KPSS

На рисунке 3.10 представлены результаты расчетного значения статистики KPSS, равный 0.111596. А также критические значения этой статистики на 1, 5 и 10%-м уровнях значимости 0.739, 0.463, 0.347 соответственно.

Поскольку расчетное значение статистики KPSS получилось меньше критического, причем на всех уровнях значимости, можно сделать вывод о том, что наш ряд стационарен.



## ***Построение авторегрессионной модели AR(p)***

При построении аддитивной модели объясняющими переменными являются функции момента времени τ. В нашем случае (для линейной модели) объясняющие переменные – лаговые переменные *yi–1, yi–2, …, yi–p*.

Модель AR(p): (где коэффициенты βj (j = 0…p) оцениваются с помощью пакета EViews, p – порядок модели)

Возмущения εi удовлетворяют условиям Гаусса–Маркова.

P1.

P2.

P3. .

При построении модели нужно решить следующие задачи:

1. определение порядка *p* авторегрессионной модели временного ряда;
2. оценивание коэффициентов βi.

*Рассмотрим решение первой задачи.*

*Ρчаст(m)* – частный коэффициент автокорреляции, соответствующий лагу *m*. Величина *Ρчаст(m)* определяется как коэффициент корреляции между двумя случайными величинами Y(),…,Y() случайных величин.

Оценкой для *Ρчаст(m)* является выборочный частный коэффициент автокорреляции *rчаст(m)*.

Предположим, что все вычисленные частные коэффициенты автокорреляции значимы до порядка *p* включительно. Тогда можно принять порядок автокорреляции модели равным *p*.

В нашем случае однозначно определить порядок p по коррелограмме не удается, поэтому будем определять порядок p подбором.

*Рассмотрим решение второй задачи.*

Наиболее часто используются следующие модели AR(p):

* авторегрессионная модель первого порядка (или модель AR(1)):

* авторегрессионная модель второго порядка (или модель AR(2)):

Коэффициентыβj (j = 0…p) будем оценивать с помощью пакета EViews.

Построим авторегрессионную модель первого порядка. Для этого необходимо в командной строке стреды EViews написать следующий код: ls y c ar(1), где y–название временного ряда. Результат выполнения указанной команды приведен на рисунке 3.11.



Рисунок 3.11 – Нахождение коэффициентов AR(p) модели

Таким образом, построенная модель AR(1) имеет следующий вид:

* + 1. ***Построение модели скользящего среднего MA(q)***

MA(q) модель имеет следующий вид:

где C, – коэффициенты модели,

q – порядок модели,

– случайные величины, образующие «белый шум». (порядок найдем перебором)

Построим модель MA(1). Коэффициенты будем считать в пакете E’Views. Для этого необходимо в командной строке стреды EViews написать следующий код: ls y c ma(1), где y–название временного ряда. Результат выполнения указанной команды приведен на рисунке 3.12.



Рисунок 3.12 – Нахождение коэффициентов MA(q) модели

Таким образом, построенная модель MA(1) имеет следующий вид:

+234088.7+0.368525

* + 1. ***Построение смешанной модели ARMA(p,q)***

Для достижения большей гибкости при построении модели ВР полезно включать в нее и авторегрессионные члены, и члены скользящего среднего.

,

где C, µj и – коэффициенты модели,

p и q – порядки модели,

– случайные величины, образующие «белый шум».

Построим модель ARMA(1,1). Коэффициенты будем считать в пакете EViews. Для этого необходимо в командной строке стреды EViews написать следующий код: ls y c ar(1) ma(1), где y–название временного ряда. Результат выполнения указанной команды приведен на рисунке 3.13.

**

Рисунок 3.13 – Построение смешанной модели ARMA(p,q)

Таким образом, построенная модель ARMA(1,1) имеет следующий вид:

* + 1. ***Сравнительный анализ построенных моделей***

Для выбора наиболее подходящей ARIMA(p,k,q) модели проведем в пакете EViews показатели адекватности для каждой модели. Во всех формулах этого пункта приняты следующие обозначения:

*n* – число фактических значений,

*m* – число коэффициентов в модели,

*yi* – фактическое наблюдение в *i*-тый момент,

 - расчётное наблюдение в *i*-тый момент.

1. Приведённый индекс детерминации (Adjusted R-squared) рассчитывается следующим образом:



где - индекс детерминации, введённый в разделе 2.3.

Данный коэффициент позволяет сравнить модели с разным числом коэффициентов. Он учитывает число коэффициентов модели, вводя штраф за дополнительные регрессоры, которые не способствуют увеличению объясняющей силы регрессии. При включении в регрессию дополнительных переменных коэффициент может уменьшаться и может быть отрицательным, если модель плохо специфицирована.

1. Стандартная ошибка регрессии (S.E. of regression) рассчитывается следующим образом:



Чем меньше значение стандартной ошибки, тем адекватнее модель.

1. Информационный критерий Акаике AIC (Akaike info criterion) рассчитывается следующим образом:



Используется для выбора лучшей модели из некоторого набора альтернативных моделей: чем меньше значение критерия, тем лучше модель.

1. Информационный критерий Шварца BIC или SC (Schwarz criterion) рассчитывается следующим образом:



Используется для выбора лучшей модели из некоторого набора альтернативных моделей – чем меньше значение критерия, тем лучше модель. Всегда выбирает лучшую модель с числом параметров, не превышающим число параметров в модели, которая была выбрана по критерию Акаике. Критерий является асимптотически состоятельным, в то время как информационный критерий Акаике смещен в сторону выбора параметризованной модели.

1. Hanna-Quinn Criterion



Используется для выбора лучшей модели из некоторого набора альтернативных моделей – чем меньше значение критерия, тем лучше модель.

В таблице 3.1 приведены расчётные значения показателей адекватности моделей AR(1), MA(1), ARMA(1,1), найденные с помощью пакета EViews.

Таблица 3.1 – Показатели адекватности моделей AR(1), MA(1), ARMA(1,1)

|  |  |  |  |
| --- | --- | --- | --- |
| Критерий | Модель | | |
| AR(1) | MA(1) | ARMA(1,1) |
| Индекс детерминации | 0,167122 | 0,159485 | 0,174280 |
| Стандартная ошибка регрессии | 556908,1 | 559455,5 | 581574,6 |
| Информационный критерий Акаике | 29,49707 | 29,50515 | 29,63222 |
| Критерий Шварца | 29,63401 | 29,64209 | 29,81480 |
| Hanna-Quinn | 29,48440 | 29,49247 | 29,61531 |

В результате получено, что из рассматриваемых моделей наилучшей является модель AR(1), так как имеет наименьшую стандартную ошибку регрессии по сравнению с моделями MA(1) и ARMA(1,1), а индекс детерминации значительно не отличается от модели ARMA(1,1).

# КОМБИНИРОВАННАЯ МОДЕЛЬ ВРЕМЕННОГО РЯДА

1. 1. **Построение комбинированной модели**

Комбинированную модель можно рассматривать как аддитивную модель временного ряда

где сезонная и периодическая составляющие отсутствуют, уравнение тренда известно, а случайная составляющая представлена моделью AR(p) вида .

Здесь

– коэффициенты модели;

p – порядок модели;

– случайные величины.

Общий вид комбинированной модели временного ряда, имеющего порядок интеграции 1:

где – расчетное значение в момент времени τi ,

– трендовая составляющая в момент времени τi ,

p – порядок авторегрессионой модели случайной составляющей ,

– коэффициенты авторегрессионой модели случайной составляющей ,

– фактические наблюдения.

Алгоритм построения комбинированной модели.

1. Провести визуальный анализ диаграммы рассеивания и убедиться в наличии тренда.
2. Построить уравнение тренда t(τ).
3. Вычислить остатки
4. Проверить гипотезу о равенстве нулю математического ожидания ряда остатков ei c помощью критерия *,*

где  *; .*

Если вычисленное значение Te непопадает в критическую область

*,*

где

то с уровнем значимости α можно считать, что математическое ожидание ряда остатков нулевое.

Для расчета можно использовать встроенную функцию Excel РАСПСТЬЮДОБР(α;n–1).

1. Вычислить коэффициенты авторегрессионной модели AR(p) ряда остатков ei с помощью пакета Eviews.
2. С учетом центрирования , построить комбинированную модель:

Рассмотрим применение алгоритма для построения комбинированной модели анализируемого временного ряда.

1. В результате визуального анализа диаграммы рассеивания, представленной на рисунке 4.1, можно сделать вывод о наличии трендовой составляющей. Кроме того, в разделе 2.1 было установлено наличие не случайной составляющей.

2–4. В разделе 2.2 было была построена полиномиальная модель трендовой составляющей

найдены значения ряда остатков ei и с помощью критерия Te показано, что ряд остатков ei имеет нулевое математическое ожидание.

1. Коэффициентов авторегрессионной модели ряда остатков будем считать в пакет E’Views. Для этого необходимо в командной строке стреды EViews написать следующий код: ls y c ar(2) ar(1), где y–название временного ряда. Результат выполнения указанной команды приведет на рисунке 4.1.



Рисунок 4.1 – Нахождение коэффициентов комбинированной модели

Согласно рисунку 4.1, искомые коэффициенты модели , ,

6.С учетом найденных коэффициентов, применим центрирование и построим комбинированную модель:



# ТОЧЕЧНЫЙ ПРОГНОЗ ПО ПОСТРОЕННЫМ МОДЕЛЯМ

Точечный прогноз заключается в получении прогнозного значения на основе построенной модели временного ряда.

## **5.1. Точечный прогноз по аддитивной модели**

Точечный прогноз по аддитивной модели заключается в получении прогнозного значения yi путем подстановки в уравнение модели момента времени τi .

В разделе 2.5 была построена аддитивная модель вида:

где значения сезонной составляющей Si приведены в таблице 5.1.

Таблица 5.1 – Значения сезонной составляющей

|  |  |  |  |
| --- | --- | --- | --- |
| 1 квартал | 2 квартал | 3 квартал | 4 квартал |
| -145394,07 | 19487,04 | -187241 | -394035,19 |

Построим точечный прогноз, для следующего момента времени, используя аддитивную модель. Поскольку последнее фактическое значение временного ряда получено для 14–го момента времени, который относится к 2–му кварталу 2018г., то построим прогноз для 15–го момента времени, который относится к 3–му кварталу:

## **5.2.** **Точечный прогноз по линейной модели**

Точечный прогноз по линейной модели заключается в получении прогнозного значения путем подстановки в уравнение модели необходимых предыдущих фактических значений временного ряда.

В главе 3 были построены несколько линейных моделей, среди которых лучшей была признана модель AR (1). Построим прогнозное значение для момента времени τn+1, следующего за моментом времени τn последнего фактического наблюдения yn. Точная формула следующая:

,

где  ­­– коэффициент модели,

 и – фактические значения временного ряда.

## **5.3. Точечный прогноз по комбинированной модели**

Точечный прогноз по аддитивной модели заключается в получении прогнозного значения путем подстановки в уравнение модели момента времени τi , а также необходимых предыдущих фактических значений временного ряда и моментов времени.

В разделе 4 разработана комбинированная модель вида:

Построим точечный прогноз, для следующего момента времени, используя комбинированную модель. Поскольку последнее фактическое значение временного ряда получено для 14–го момента времени, то построим прогноз для 15–го момента времени, полагая   
τi =15, τi–1=14,τi–2=13, yi–1=416285, yi–2=385878:

**5.4. Сравнительный анализ построенных прогнозов**

Таблица 5.2 – Сравнительный анализ построенных прогнозов

|  |  |  |  |
| --- | --- | --- | --- |
| Характеристики | Модель | | |
| Аддитивная | Линейная | Комбинированная |
| Приведённый индекс детерминации R2adj | 0,13974 | 0,01569 | 0,0641 |
| Прогноз для момента времени, следующего за моментом последнего фактического наблюдения, yn+1 | -1432712 | 371846,69 | |  | | --- | | -841866 | |

# ЗАКЛЮЧЕНИЕ

В ходе работы проанализирована диаграмма рассеяния финансового состояния кредитной организации ПАО «СКБ–БАНК», выявлено наличие тренда, на основании чего для анализа и дальнейшего прогнозирования построены следующие модели:

* аддитивная модель, включающая трендовую и сезонную составляющие;
* модели ряда первых разностей AR(1), MA(1), ARMA(1,1);
* комбинированная модель.

Установлено, что из построенных моделей лучшей является аддитивная модель, тогда прогнозное значение финансового состояния на 3 квартал 2018 г. составляет -1432712 миллиардов рублей, соответственно.

# 

# СПИСОК ЛИТЕРАТУРЫ

1. Шанченко, Н. И. Лекции по эконометрике: учебное пособие для студентов высших учебных заведений, обучающихся по специальности «Прикладная информатика (в экономике)» / Н. И. Шанченко. – Ульяновск: УлГТУ, 2008. – 139 с.
2. Э.Е. Тихонов. Методы прогнозирования в условиях рынка: учебное пособие. – Невинномысск, 2006. – 221 с.
3. Официальный сайт ПАО «СКБ–БАНК», доступен по ссылке: [https://www.sberbank.com](https://www.sberbank.com/ru/about) (дата обращения: 01.10.2019г.).

# ПРИЛОЖЕНИЕ А

Таблица А.1 – Значение фактических наблюдений финансового состояния ПАО «СКБ–БАНК»

|  |  |  |
| --- | --- | --- |
| Номер момента времени (i) | Период | Фактическое наблюдение () |
| 1 | 2 кв. 2018 | 416285 |
| 2 | 1 кв. 2018 | 385878 |
| 3 | 4 кв. 2017 | 1400376 |
| 4 | 3 кв. 2017 | 457284 |
| 5 | 2 кв. 2017 | 416727 |
| 6 | 1 кв. 2017 | 92097 |
| 7 | 4 кв. 2016 | -716318 |
| 8 | 3 кв. 2016 | -1174833 |
| 9 | 2 кв. 2016 | 406003 |
| 10 | 1 кв. 2016 | 119011 |
| 11 | 4 кв. 2015 | 1017501 |
| 12 | 3 кв. 2015 | 845311 |
| 13 | 2 кв. 2015 | 211148 |
| 14 | 1 кв. 2015 | 345155 |