

Построение прогностической модели предоставления услуги при подключении и отключении к Internet

Перепадина Екатерина Николаевна

*Приамурский государственный университет имени Шолом-Алейхема
студент*

Баженов Руслан Иванович

*Приамурский государственный университет имени Шолом-Алейхема
к.п.н., доцент, зав. кафедрой информационных систем, математики и
методики обучения*

Аннотация

В статье проведен анализ статистических данных предприятия и составлен краткосрочный прогноз о количестве подключенных и отключенных абонентов в среде STATISTICA.

Ключевые слова: прогностическая модель, абоненты Internet, STATISTICA.

The construction of prognostic models of providing services by connecting and disconnecting to the Internet

Perepadia Ekaterina Nikolaevna

*Sholom-Aleichem Priamursky State University
student*

Bazhenov Ruslan Ivanovich

*Sholom-Aleichem Priamursky State University
Candidate of pedagogical sciences, associate professor, Head of the Department
of Information Systems, Mathematics and teaching methods*

Abstract

In the article the analysis of statistical data of the company and compiled short-term forecast on the number of connected and disconnected subscribers in STATISTICA.

Keywords: predictive model, the subscribers Internet, STATISTICA.

В настоящее время рынок предоставления услуг по подключению к информационно-телекоммуникационной сети «Интернет» является достаточно развитым, в том числе на территории Еврейской автономной области.

В условиях жесткой конкурентной борьбы на фоне экономического кризиса для каждого участника рынка важное значение имеет прогнозирование объемов оказываемых услуг. Эффективный анализ

статистических показателей, характеризующих результаты деятельности предприятия, позволяет определить динамику рынка, своевременно реагировать на изменение его конъюнктуры, проводить гибкую финансовую и кадровую политику, минимизируя производственные издержки. Именно поэтому построение прогностических моделей актуально и жизненно необходимо для каждого предприятия, оказывающего данные услуги.

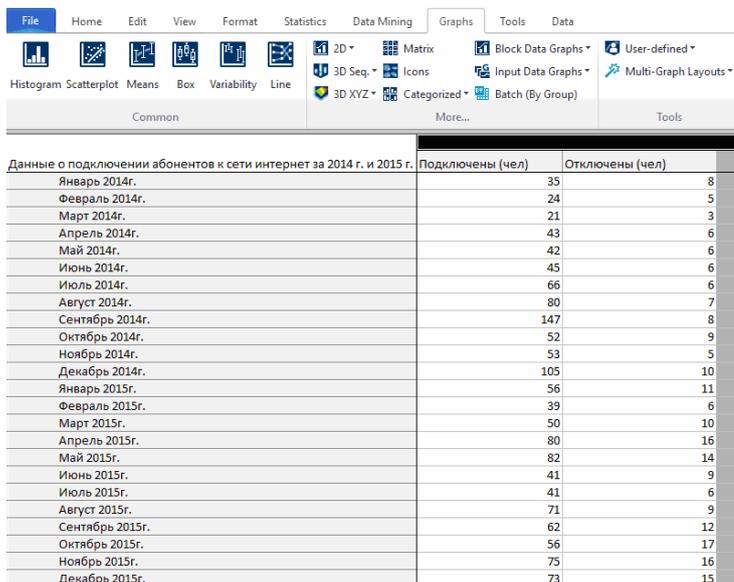
О статистическом анализе данных в своих работах пишут В.П. Боровиков, И.П.Боровиков, а также А.А.Халафян, о возможностях инструмента dataminer в своей статье описывает П.С. Большаков. [1,2,3]. Для программы Statistica разработчиками написано полное электронное руководство [4, 5]. Р.И.Баженов и др. изучали возможности применения различных видов анализа [6-14]. Особенности моделирования латентных изменений с помощью AMOS SPSS рассмотрел Р.И.Остапенко [15, 16]. С.А.Суслов исследовал проблемы анализа в сельском хозяйстве [17, 18]. Скоринг экспериментальных данных с применением прогностических моделей в среде пакета SPSS показал А.А.Серов [19,20]. Зарубежные ученые так же исследовали проблемы применения различных методов анализа [21, 22].

Цель данной статьи построение прогностической модели предоставления услуг подключения к сети Интернет на основе статистических данных предприятия.

В настоящее время разработано достаточно большое количество статистических программ, способных выполнять широкий круг задач. При проведении настоящего исследования для построения прогностической модели предоставления услуг подключения к сети Интернет мы считаем целесообразным применить программный пакет STATISTICA.

В качестве исходных данных были взяты данные о количестве подключенных и отключенных абонентов к сети Интернет предприятия за 24 месяца.

Данные в программе STATISTICA представлены на рис. 1



Данные о подключении абонентов к сети интернет за 2014 г. и 2015 г.	Подключены (чел)	Отключены (чел)
Январь 2014г.	35	8
Февраль 2014г.	24	5
Март 2014г.	21	3
Апрель 2014г.	43	6
Май 2014г.	42	6
Июнь 2014г.	45	6
Июль 2014г.	66	6
Август 2014г.	80	7
Сентябрь 2014г.	147	8
Октябрь 2014г.	52	9
Ноябрь 2014г.	53	5
Декабрь 2014г.	105	10
Январь 2015г.	56	11
Февраль 2015г.	39	6
Март 2015г.	50	10
Апрель 2015г.	80	16
Май 2015г.	82	14
Июнь 2015г.	41	9
Июль 2015г.	41	6
Август 2015г.	71	9
Сентябрь 2015г.	62	12
Октябрь 2015г.	56	17
Ноябрь 2015г.	75	16
Декабрь 2015г.	73	15

Рисунок 1 – Исходные данные

В первую очередь произведем анализ данных о подключении.

Для начала необходимо построить линейный график заданной переменной. (Graphs => 2D Graphs => Line Plots (Variables))

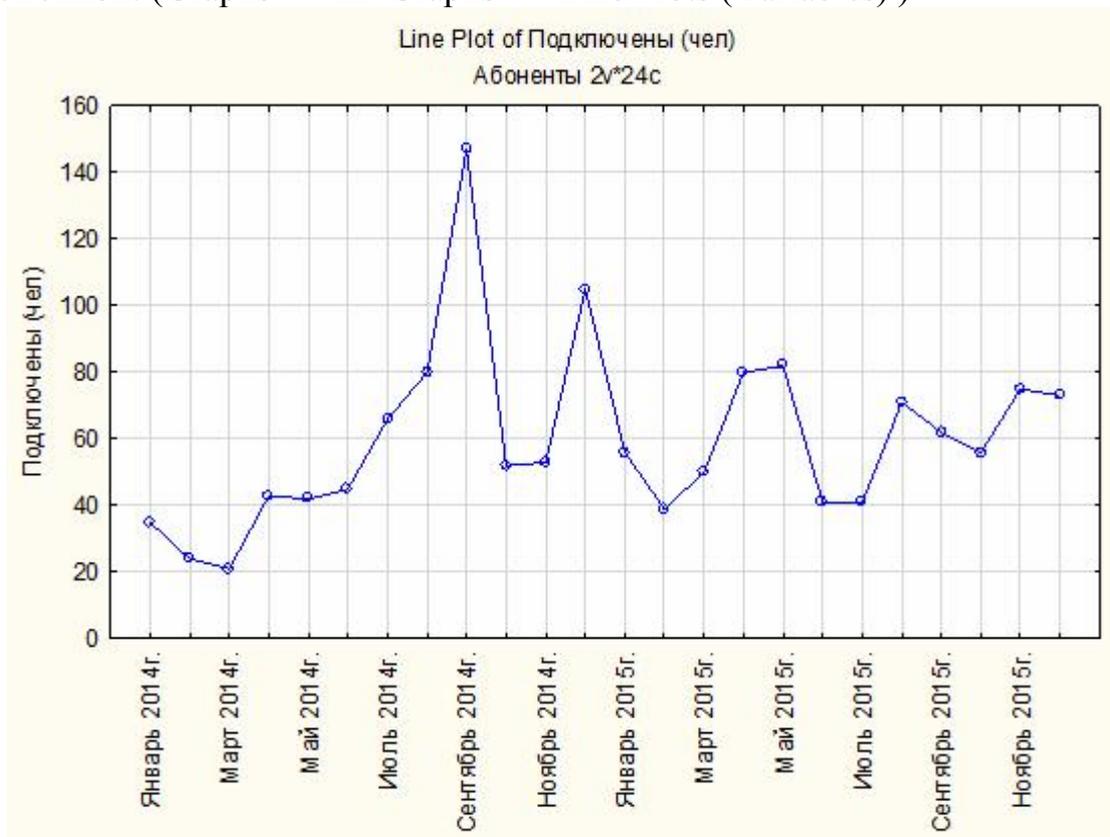


Рисунок 2 –График исходных данных

По представленному графику достаточно сложно определить, какие факторы оказывают влияние на формирование временного ряда (в силу незначительного количества значений ряда).

Также на графике можно заметить, что выбросы - скачки количества подключенных к сети, пусть и не очень большие, но имеются. Однако, этот факт значительно на составление прогноза не повлияет.

Для более детального ознакомления с рядом можно построить график автокорреляции (Statistics =>AdvancedModels =>TimeSeries/Forecasting =>ARIMA&autocorrelationfunctions =>Autocorrelations), который имеет следующий вид:

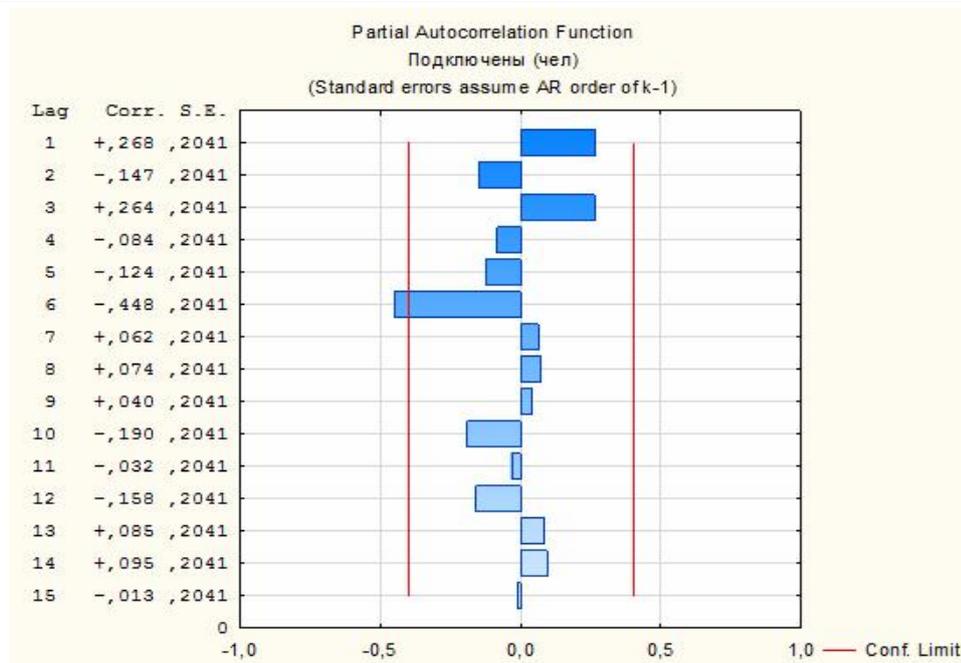


Рисунок 3 – График автокорреляции

Автокорреляция не является медленно убывающей, что говорит об отсутствии тесной взаимосвязи между соседними элементами временного ряда, и как следствие можно говорить об отсутствии тренда в этом ряде. Также из вида автокорреляционной функции можно говорить об отсутствии сезонности (нет явной зависимости между пиками на каких-либо лагах).

Можем перейти к построению модели. Для прогнозирования использовалась модель ARIMA. (Statistics => Advanced Models => Time Series/Forecasting => ARIMA & autocorrelation functions) Параметры построения модели отображены на рис. 4.

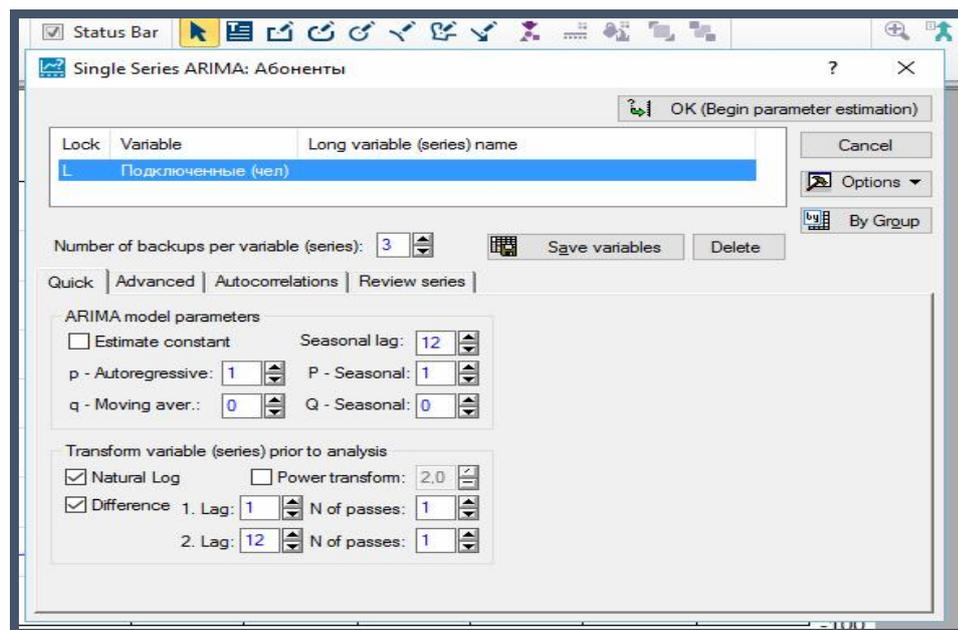


Рисунок 4 – Параметры прогноза

Результат краткосрочного прогнозирования отображен на рис.5.

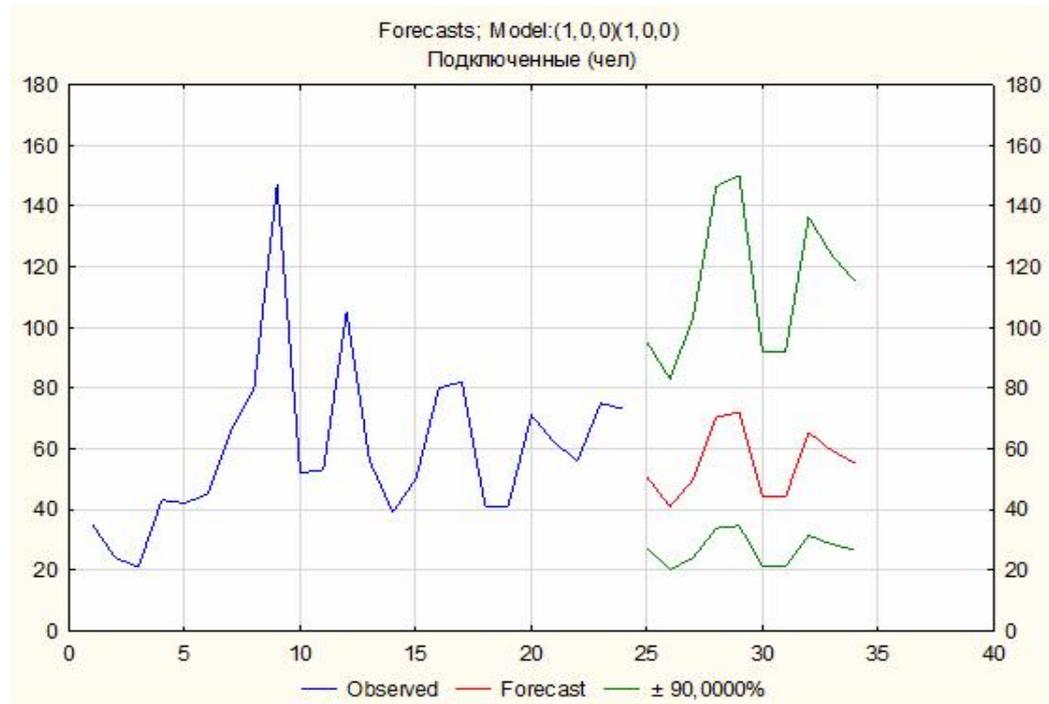


Рисунок 5 – График краткосрочного прогноза

По результатам прогнозирования количество подключившихся в следующем месяце – январе 2016 г. составит - 50,67229 человек.

Можно выполнить проверку, рассмотрев график нормального распределения остатков. (Statistics =>AdvancedModels =>TimeSeries/Forecasting =>ARIMA&autocorrelationfunctions =>Distribushionofresiduals). Если остатки нормально распределены, то все точки должны группироваться около прямой линии на графике (рис.6).

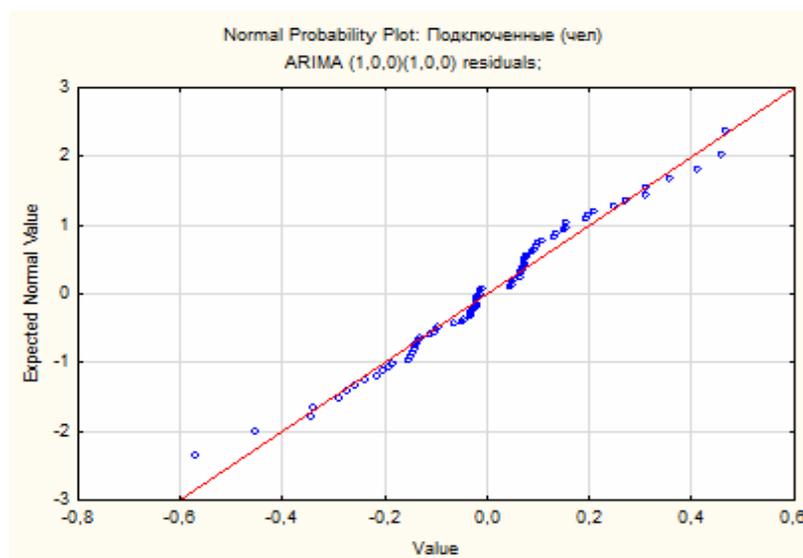


Рисунок 6 – График нормального распределения остатков

Как и в предыдущем случае, в первую очередь произведем анализ данных о подключении. Построить линейный график заданной переменной.

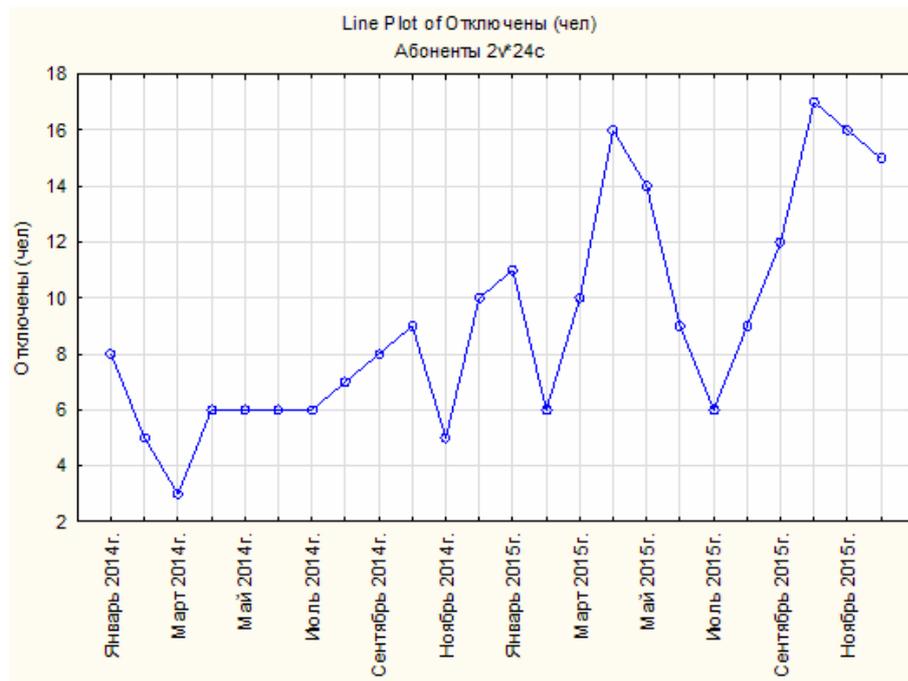


Рисунок 7 – График исходных данных

Как и в предыдущем случае, по представленному графику достаточно сложно определить, какие факторы оказывают влияние на формирование временного ряда (в силу незначительного количества значений ряда).

Для более детального ознакомления с рядом можно построить график автокорреляции (Statistics =>AdvancedModels =>TimeSeries/Forecasting =>ARIMA&autocorrelationfunctions =>Autocorrelations), который имеет следующий вид:

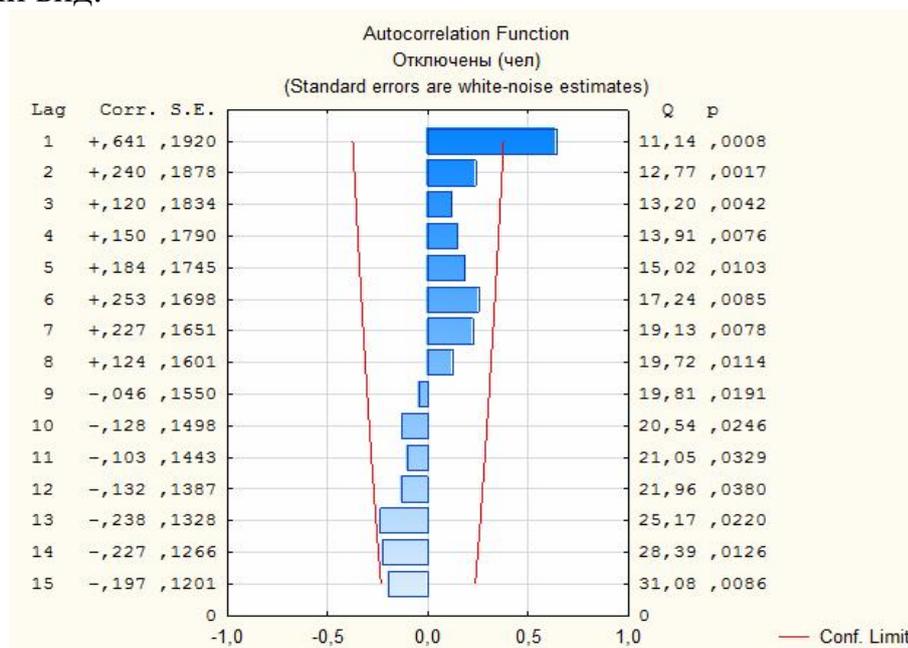


Рисунок 8 – График автокорреляции

Из графика автокорреляционной функции также сложно установить присутствие тренда (тесной взаимосвязи между соседними элементами временного ряда). Также из вида автокорреляционной функции можно говорить об отсутствии сезонности (нет явной зависимости между пиками на каких-либо лагах). Скорее всего это связано с недостаточным количеством данных во временном ряде.

Однако можно выполнить прогнозирование. Для построения прогноза также используется модель ARIMA. (Statistics => Advanced Models => Time Series/Forecasting => ARIMA & autocorrelation functions). Параметры построения модели отображены на рис. 9.

Рисунок 9 – Параметры прогноза

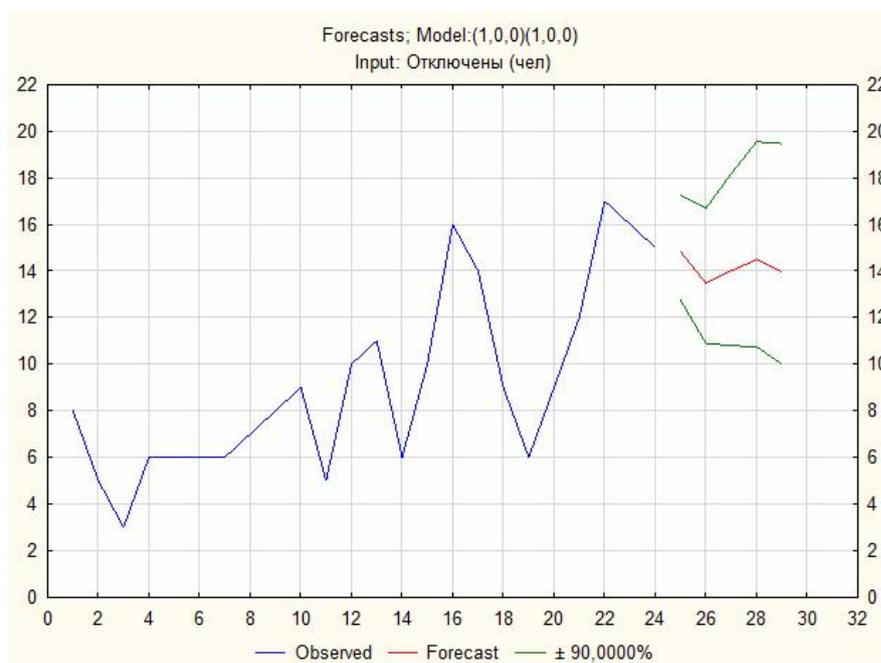


Рисунок 10 – График краткосрочного прогноза

По результатам прогнозирования (рис.10) количество отключившихся в следующем месяце – январе 2016 г. составит - 15,89677 человек.

В результате данного исследования предполагаемое количество подключенных абонентов составит 50,67229 человек, а отключенных 15,89677 человек, что обеспечит небольшой, но стабильный прирост абонентов, это позволит предприятию получить прибыль и сохранить свои позиции на рынке.

Библиографический список

1. Боровиков В.П., Боровиков И.П., STATISTICA. Статистический анализ и обработка данных в среде Windows. Издание 2 изд. М.: Информационно-издательский дом «Филинь», 1998. 608 с.
2. Халафян А.А. STATISTICA 6. Статистический анализ данных. М.: Бином, 2007. 512с.
3. Большаков П.С. Возможности Statistica Data Miner. // URL: <http://images.nature.web.ru/nature/2003/04/11/0001193684/04.pdf>
4. STATISTICA (Версия 6.1). Электронное руководство.
5. StatSoft. // URL: <http://www.statsoft.ru/>
6. Николаев С.В., Пронина О. Ю., Баженов Р.И. Исследование методов интеллектуального анализа для формирования краткосрочного прогноза в программной среде STATISTICA // Экономика и менеджмент инновационных технологий. 2015. №№7(46). С. 89-102.
7. Иванченко М.А., Кельнер Т.А., Баженов Р.И. Применение ABC-анализа для исследования ассортимента аптечных препаратов // Постулат. 2016. № 4. С. 24.
8. Баженов Р.И., Векслер В.А., Гринкруг Л.С. RFM-анализ клиентской базы в прикладном решении 1С:Предприятие 8.3 // Информатизация и связь. 2014. № 2. С. 51-54.
9. Остроушко А.А., Баженов Р.И. Анализ ассортимента электротоваров с использованием abc-анализа // Экономика и менеджмент инновационных технологий. 2014. № 10 (37). С. 73-81.
10. Бронштейн К.С., Наумов А.А., Баженов Р.И. Применение классического abc-анализа для анализа ассортимента блюд кафе // Экономика и менеджмент инновационных технологий. 2014. № 11 (38). С. 100-110.
11. Пронина О.Ю., Баженов Р.И. Исследование методов регрессионного анализа программной среды EVIEWS // Nauka-Rastudent.ru. 2015. № 1 (13). С. 45.
12. Широкова Н.А., Баженов Р.И. Применение корреляционного анализа для исследования данных спортивных показателей студентов в среде SPSS // Nauka-Rastudent.ru. 2015. № 6 (18). С. 25.
13. Лагунова А.А., Баженов Р.И. Разработка в среде gretl регрессионной модели рынка вторичного жилья г. Биробиджана // Nauka-Rastudent.ru. 2015. № 1 (13). С. 40.

14. Муллинов Д.О., Баженов Р.И. Разработка в среде eviews регрессионной модели рынка гаражных помещений г. Биробиджана // Nauka-Rastudent.ru. 2015. № 1 (13). С. 43.
15. Остапенко Р.И. Особенности моделирования латентных изменений с помощью AMOS SPSS // Перспективы науки и образования. 2014. № 1 (7). С. 89-95.
16. Остапенко Р.И. Особенности анализа лонгитюдных данных в психолого-педагогических исследованиях с помощью AMOS SPSS // Современные научные исследования и инновации.. 2014. №2014. № 4 (36). С. 74.
17. Суслов С.А. Кластерный анализ: сущность, преимущества и недостатки // Вестник НГИЭИ. 2010. Т. 1. № 1. С. 51-57.
18. Суслов С.А. Применение динамического программирования в отраслях сельского хозяйства // Экономический анализ: теория и практика. 2011. № 31. С. 56-60.
19. Серов А.А. Скоринг экспериментальных данных с применением прогнозных моделей в среде пакета SPSS // Традиции и новации в профессиональной подготовке и деятельности педагога. Материалы Всероссийской научно-практической конференции преподавателей и студентов. Ответственный редактор В.П.Анисимов. Тверь: Тверской государственный университет, 2013. С. 53-54.
20. Серов А.А. Некоторые вопросы применения ресамплинговых методов в эконометрическом анализе // В сборнике: Математика, статистика и информационные технологии в экономике, управлении и образовании Ответственный редактор А.А. Васильев. 2015. С. 211-216
21. Miner G. Practical text mining and statistical analysis for non-structured text data applications. Academic Press, 2012.
22. Miner G., Nisbet R., Elder IV J. Handbook of statistical analysis and data mining applications. Academic Press, 2009.