

УДК 004.4:519.23

ПРОГНОЗИРОВАНИЕ НА ОСНОВЕ МУЛЬТИПЛИКАТИВНОЙ И МНОЖЕСТВЕННОЙ РЕГРЕССИОННОЙ МОДЕЛИ СЕЗОННОСТИ**Сальников В.И.***Санкт-Петербургский имени В.Б. Бобкова филиал Российской таможенной академии***FORECASTING BASED ON MULTIPLE REGRESSION AND MULTIPLICATIVE SEASONALITY MODEL****Salnikov V.I.***St. Petersburg named after V.B. Bobkova branch of the Russian Customs Academy***Аннотация**

В статье представлена методика проведения статистического исследования динамики и сезонности развития экономического явления с применением электронных таблиц.

Ключевые слова: статистика, сезонность, множественная регрессионная модель, электронные таблицы.

Abstract

The article presents a method of statistical study of the dynamics and seasonality of the economic phenomenon with the use of spreadsheets.

Keywords: statistics, seasonality, multiple regression model, spreadsheets.

При проведении статистического анализа нередко приходится сталкиваться с исследованием временных рядов, содержащих периодические (сезонные) колебания. Для прогнозирования на основе указанных временных рядов необходим анализ сезонной компоненты и корректировка трендового прогноза на величину такой компоненты. При анализе используются, как правило, квартальные данные, но могут быть использованы данные и за другие промежутки времени, например, месяцы.

При проведении анализа сезонности и прогнозирования целесообразно пользоваться возможностями информационных технологий, в частности электронных таблиц. В приложении Microsoft Excel имеются функции, которые ускоряют статистические расчеты и позволяют проводить их без использования статистических формул. Кроме того, электронные таблицы позволяют обрабатывать большой объем данных, применять копирование формул и непосредственно наблюдать графическое представление изучаемого явления. [1, 2]

Обычно для прогнозирования экономических показателей составляется уравнение развития явления, и оцениваются его параметры. Вначале исходные данные вносятся в таблицу в хронологическом порядке. Далее выполняется аппроксимация

(сглаживание) и подбирается функция (линия тренда), отражающая тенденцию развития. Аппроксимацию можно произвести с помощью встроенных функций Excel, а также путем построения диаграммы и линии тренда. Затем проводится исследование влияния сезонной компоненты на изучаемое явление. После этого можно вычислить прогнозное значение изучаемого показателя на последующие периоды.

Для исследования сезонных колебаний могут использоваться две модели: мультипликативная и множественная регрессионная (аддитивная). При использовании мультипликативной модели сначала определяется основная тенденция, а затем определяются индексы сезонной волны. Прогноз по тренду корректируется на соответствующий прогнозному кварталу индекс сезонной волны. При использовании множественной регрессионной модели сразу строится уравнение, учитывающее не только основную тенденцию, но и влияние каждого сезона. Поэтому при прогнозировании достаточно только подставить в уравнение прогнозные значения факторов.

При выявлении тенденции часто используются линейные функции. Для построения модели вида $y = b + m \cdot x$ и $y = b + m_1 \cdot x_1 + m_2 \cdot x_2 + \dots + m_n \cdot x_n$ при наличии двух массивов данных признаков в Excel исполь-

зуется функция ЛИНЕЙН. В качестве аргументов функции вводятся диапазоны данных, содержащие значения признака-результата и признака-фактора (факторов). При анализе динамики и сезонности с использованием мультипликативной модели признаком-фактором является фактор времени, который вводится в порядке возрастания с равномерным шагом от единицы. При анализе сезонности с использованием множественной регрессионной модели признаками-факторами являются как фактор времени, так и дополнительные факторы трех сезонов, которые принимают значения 0 или 1. В этом случае все признаки-факторы служат единым аргументом для функции.

Функция ЛИНЕЙН возвращает массив, содержащий параметры уравнения тенденции в порядке $\{m_1, \dots, m_n, b\}$. Она может использоваться и для вычисления параметров других трендов, например, логарифмического или полиномиального, которые также часто используются при анализе сезонности. В этом случае рассчитываются дополнительные параметры, например, логарифм от фактора времени или квадрат фактора времени. Для построения экспоненциальной модели вида $y = b * t^x$ используется функция ЛГРФПРИБЛ, которая работает аналогично функции ЛИНЕЙН и также возвращает массив, содержащий параметры уравнения в порядке $\{m, b\}$. Однако она редко используется при анализе сезонности, поскольку предполагает значительный рост изучаемого показателя в каждом периоде времени.

Для осуществления выбора наилучшего тренда для прогнозирования необходимо вычисление коэффициента детерминации. Для этого в функциях ЛИНЕЙН и ЛГРФПРИБЛ предусмотрена возможность вывода дополнительной регрессионной статистики. При записи функции значение ее последнего аргумента необходимо задать как ИСТИНА. В этом случае коэффициент детерминации будет выведен в третьей строке первого столбца таблицы-результата. Чем ближе данный коэффициент к единице, тем лучше подходит данный тренд для прогнозирования.

Другим способом определения уравнения является построение точечной диаграммы в виде зависимости признака-результата от признака-фактора. Строится линия тренда путем выбора соответствующего пункта в контекстном меню. В появив-

шемся окне необходимо выбрать подходящий тип линии тренда, уравнение на диаграмме и поместить на диаграмму величину достоверности аппроксимации, для чего устанавливаются соответствующие флажки. Выбор типа линии тренда основывается на величине достоверности аппроксимации, которая должна быть максимальной среди всех предложенных трендов. Однако данный способ имеет недостатки. Необходимо устанавливать больше десятичных знаков в уравнении, чтобы прогноз был более точен. Требуется переписать уравнение с параметрами в виде констант в ячейку, и в случае обновления данных обновить параметры тренда. Построить множественную регрессионную модель на диаграмме и получить ее параметры невозможно.

Рассмотрим алгоритмы проведения расчетов при выявлении сезонной компоненты по двум предложенным моделям более подробно.

При использовании мультипликативной модели необходимо выполнить расчет индексов сезонной волны. Сначала необходимо выявить тенденцию явления и построить уравнение тренда. Определив параметры тренда, следует рассчитать теоретические уровни временного ряда, подставляя фактор времени в уравнение. Далее необходимо определить индексы колеблемости, которые получаются отношением фактических уровней ряда и теоретических. Затем необходимо осреднить индексы колеблемости по одноименным кварталам за все годы. В результате получаются индексы сезонности, сумма которых должна быть равна 4. Если сумма всех индексов сезонности не равна 4, то рассчитывается поправочный коэффициент путем деления 4 на сумму индексов сезонности. Затем на поправочный коэффициент умножаются все индексы сезонности, и в результате получаются индексы сезонной волны. Зная индексы сезонной волны, можно рассчитать прогнозное значение изучаемого показателя на определенный квартал. Трендовый прогноз рассчитывается по уравнению, предварительно выписанному в ячейку в виде формулы, и умножается на индекс сезонной волны соответствующего квартала. Также для линейного тренда можно использовать функцию ТЕНДЕНЦИЯ, для экспоненциального – функцию РОСТ с последующим умножением на индекс сезонной волны соответствующего квартала.

Таблица 1

Анализ сезонности импорта продовольственного товара с использованием мультипликативной модели

Год	Квартал	x1	x2	Количество, тыс. т (факт)	Количество, тыс. т (тренд)	Индекс колеблемости	Количество, тыс. т (теор)
1	1	1	1	124,35	153,44	0,810	129,08
	2	4	2	194,14	150,61	1,289	186,52
	3	9	3	162,06	148,01	1,095	154,85
	4	16	4	125,20	145,63	0,860	127,31
2	1	25	5	115,28	143,47	0,803	120,70
	2	36	6	169,27	141,54	1,196	175,28
	3	49	7	141,76	139,82	1,014	146,29
	4	64	8	116,70	138,33	0,844	120,93
3	1	81	9	124,71	137,07	0,910	115,31
	2	100	10	167,33	136,02	1,230	168,44
	3	121	11	139,24	135,20	1,030	141,45
	4	144	12	123,71	134,60	0,919	117,66
	1	169	13		134,22		112,91
	2	196	14		134,06		166,02

Параметры тренда			Индексы сезонной волны		
m2	m1	b	i1	0,841	0,841
-3,159	0,111	156,488	i2	1,238	1,238
			i3	1,046	1,046
			i4	0,874	0,874
			Сумма	3,99998	4,0000
			Попр.коэф.	1,00001	

При использовании множественной регрессионной модели параметры уравнения вычисляются с использованием функции ЛИНЕЙН. В качестве факторных признаков используются помимо фактора времени три дополнительных фактора сезонности. В первом квартале значения факторов сезонности принимаются равными нулю, в последующих значение одного из факторов равно единице, а остальные – нулю. Таким способом, формируется таблица факторных признаков. После вычисления параметров множественной регрессии необходимо рассчитать теоретические значения исследуемого

показателя для контроля правильности проведенных расчетов. Затем следует рассчитать прогноз, подставив в уравнение значения факторных признаков для требуемого квартала. Уравнение при этом также предварительно выписывается в ячейку в виде формулы.

Рассмотрим пример прогнозирования с использованием мультипликативной модели. Имеются квартальные данные о физических объемах импорта сезонного продовольственного товара за три прошедших года.

Допустим, что требуется выполнить прогноз на первые два квартала текущего

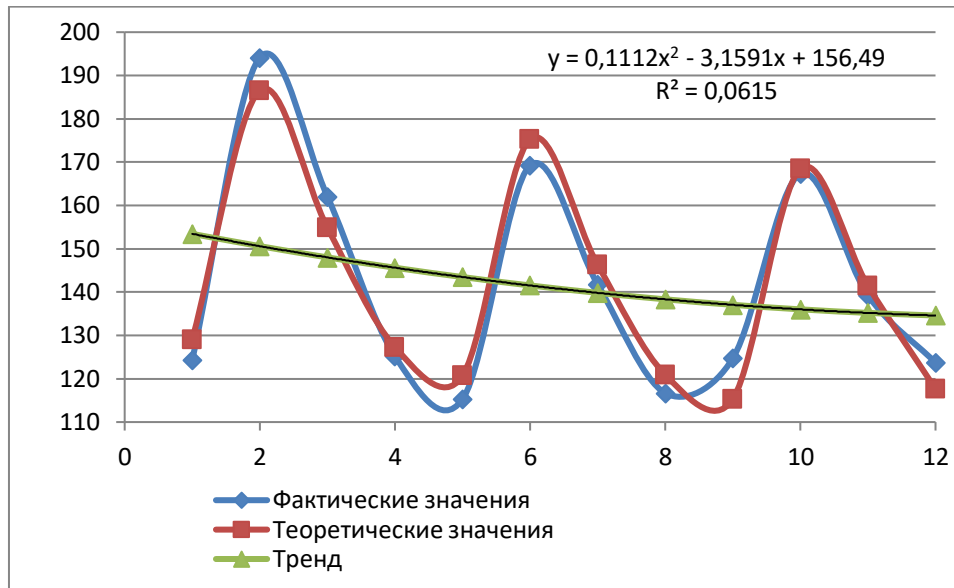


Рис 1. Фактические, трендовые и теоретические значения импорта товара

Таблица 2

Анализ сезонности импорта продовольственного товара с использованием множественной регрессионной модели

Год	Квартал	x_1^2	x_1	x_2	x_3	x_4	Количество, тыс.т (факт)	Количество, тыс. т (теор)
1	1	1	1	0	0	0	124,35	132,93
	2	4	2	1	0	0	194,14	186,03
	3	9	3	0	1	0	162,06	154,43
	4	16	4	0	0	1	125,20	126,24
2	1	25	5	0	0	0	115,28	118,28
	2	36	6	1	0	0	169,27	173,75
	3	49	7	0	1	0	141,76	144,52
	4	64	8	0	0	1	116,70	118,71
3	1	81	9	0	0	0	124,71	113,12
	2	100	10	1	0	0	167,33	170,96
	3	121	11	0	1	0	139,24	144,11
	4	144	12	0	0	1	123,71	120,67
	1	169	13	0	0		117,46	
	2	196	14	1	0		177,67	

Параметры тренда					
m5	m4	m3	m2	m1	b
5,19	30,01	57,65	-5,44	0,297	138,077

года. Для выполнения задачи определены параметры тренда с помощью функции ЛИНЕЙН, рассчитаны значения импорта по

тренду, индексы колеблемости, сезонности и сезонной волны, затем рассчитаны теоретические значения импорта по

тические значения импорта. Автоматизировать расчет индексов сезонности можно с помощью функции СУММЕСЛИ. Эту функцию можно применять для расчета сумм индексов колеблемости, а в качестве критерия использовать номер квартала. Для вычисления теоретических значений показателя, скорректированных на индекс сезонной волны, можно использовать функцию ВПР, позволяющую подставлять в формулу значение индекса сезонной волны необходимого квартала. При использовании функции ВПР расчет ускоряется.

Далее рассмотрим пример прогнозирования с использованием множественной регрессионной модели. Статистика, выводимая функцией ЛИНЕЙН, показывает, что наилучшим трендом для прогнозирования также является полиномиальный. Для полиномиального тренда коэффициент детерминации составляет 0,9408, а для линейного 0,9244, для логарифмического 0,9296. Для построения данной модели требуется пять переменных. Первая переменная – это квадрат фактора времени, вторая – сам фактор времени, при условии, что первый

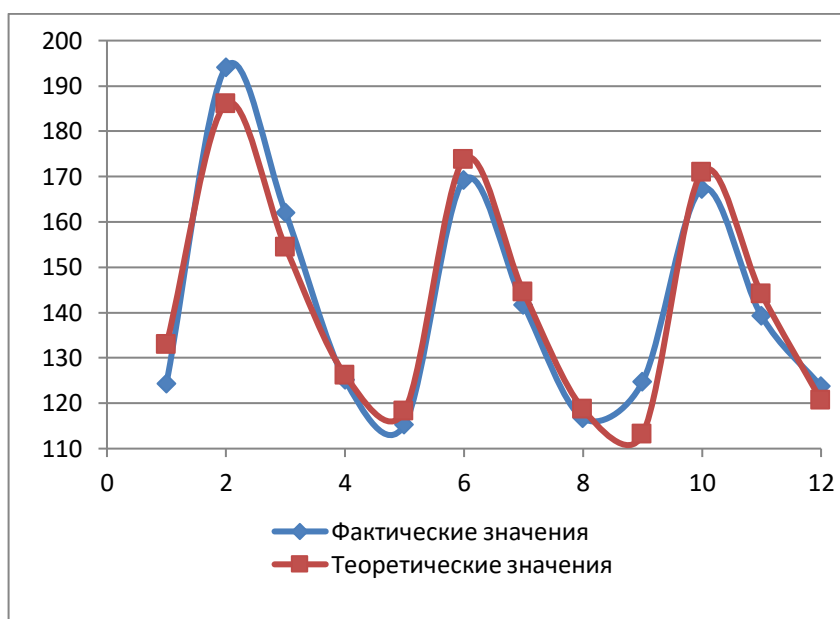


Рис. 2. Фактические и теоретические значения импорта товара

Построение диаграммы и линии тренда показали, что наилучшим трендом для анализа является полиномиальный. Для вычисления его параметров дополнительно рассчитывается квадрат фактора времени, который должен стоять в таблице перед фактором времени. Результаты представлены ниже (табл. 1).

Фактические, трендовые и теоретические значения импорта представлены ниже (рис. 1). Видно, что импорт имеет тенденцию к снижению. Модель достаточно хорошо описывает изучаемое явление. Во втором и третьем квартале наблюдается рост поставок, в первом и четвертом – снижение поставок. Прогноз составляет 112,91 тыс.т на первый квартал и 166,02 тыс.т на второй квартал текущего года.

квартал принят за базовый уровень. Остальные переменные – это коэффициенты-индикаторы сезонности, которые принимают значения 0 или 1. В первом квартале эти переменные равны нулю, а в последующих кварталах одна из переменных равна единице, а остальные – нулю. После ввода всех переменных вычисление параметров тренда производится с помощью функции ЛИНЕЙН, при этом получают 6 параметров. На их основе рассчитываются теоретические значения импорта, далее строится диаграмма. Данные сведем в таблицу (табл. 2).

Фактические и теоретические значения импорта представлены ниже (рис.2).

Уравнение для данной модели будет иметь вид:

$$y = 138,077 + 0,297x_{12} - 5,44x_1 + 57,65x_2 + 30,01x_3 + 5,19x_4.$$

По рисунку видно, что модель достаточно хорошо описывает изучаемое явление.

Во втором и третьем квартале наблюдается существенный рост поставок относительно первого квартала. В четвертом квартале наблюдается незначительный рост поставок относительно первого квартала, но существенное снижение относительно второго и третьего кварталов. Сама динамика импорта также имеет тенденцию к снижению, поэтому в первом и в четвертом кварталах в итоге получаются близкие значения в пределах года. Прогноз составляет 117,46 тыс.т на первый квартал и 177,67 тыс.т на второй квартал текущего года. Прогноз можно рассчитать, подставляя значения признаков-факторов в уравнение, а также с использованием функции =ТЕНДЕНЦИЯ. Результаты получаются одинаковые.

Таким образом, модели дают схожие результаты при прогнозировании. Множественная регрессионная модель при этом позволяет оценить рост или снижение в абсолютном выражении. Для ее реализации любой квартал можно брать за базу сравнения, а результаты прогнозирования при этом будут одинаковы. Меняться будут только параметры уравнения, коэффициенты при факторах сезонности в них будут отражать рост или снижение показателя относительно базового квартала.

Предложенная в статье методика может быть использована в образовательном процессе при проведении практических занятий по статистике с применением программных средств, а также при проведении различных статистических исследований в таможенных органах.

Список литературы

1. Сальников В.И. Методика формирования общих профессиональных компетенций студентов на примере алгоритмизации расчета таможенных платежей // Ученые записки Санкт-Петербургского имени В.Б.Бобкова филиала РТА. – 2016. – № 4 (60). – С. 109-117.

2. Сальников В.И. Построение систем принятия решений на основе операций алгебры логики // Ученые записки Санкт-Петербургского имени В.Б. Бобкова филиала Российской таможенной академии. 2019. № 2 (70). С. 101–107..

Поступила в редакцию 09.08.2019

Сведения об авторе:

Сальников Владислав Игоревич – старший преподаватель кафедры таможенного дела Санкт-Петербургского филиала Российской таможенной академии, e-mail: visrta@yandex.ru

Электронный научно-практический журнал "Бюллетень инновационных технологий" (ISSN 2520-2839) является сетевым средством массовой информации регистрационный номер Эл № ФС77-73203 по вопросам публикации в Журнале обращайтесь по адресу bitjournal@yandex.ru