МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра информационных систем и технологий

УТВЕРЖДАЮ
Проректор по учебной работе
__________ О. Г. Локтионова
«25» 01 __________ 2018г.

МЕТОДЫ ПРОГНОЗИРОВАНИЯ

Методические указания по выполнению
лабораторных работ по дисциплине
«Компьютерные методы прогнозирования и планирования»
для студентов направления подготовки
09.03.03 Прикладная информатика

Курск 2018

УДК 004.82 (075.8)

Составитель: Т.И.Лапина

Рецензент

Доктор технических наук, профессор Р.А.Томакова

Методы прогнозирования: методические указания по выполнению лабораторных работ / Юго-Зап. гос. ун-т; сост.: Т. И. Лапина, Курск, 2018. 45с.: ил.26, табл. 5, Библиогр.: с. 5.

Содержат краткие теоретические сведения о способах прогнозирования и планирования на предприятии, а также об инструментальных средах используемых при составлении прогнозов.

Методические указания соответствуют требованиям программ по направлениям подготовки бакалавров: 09.03.03 Прикладная информатика.

Предназначены для студентов направления подготовки бакалавров 09.03.03 Прикладная информатика дневной и заочной форм обучения.

Текст печатается в авторской редакции

Подписано в печать Формат 60х84 1/16.

Усл. печ. л. . Уч. – изд. л. .Тираж 100 экз. Заказ. Бесплатно.

Юго - Западный государственный университет.

305040, г. Курск, ул. 50 лет Октября, 94.

Лабораторная работа № 1

Определение динамики временных рядов экспериментальных данных

Цель работы: ознакомиться с методами анализа динамики и сглаживания данных и научиться проводить э сглаживание методами скользящей средней и экспоненциального сглаживания в программе Excel.

Общие положения

Прогноз динамики ряда с помощью среднего абсолютного прироста соответствует его представлению в виде прямой, проведенной через две крайние точки. В этом случае, чтобы получить прогноз на L шагов вперед
(L- период упреждения), достаточно воспользоваться следующей формулой:

$$\hat{Y}_{n+T} = y_n + T \cdot \overline{\Delta Y} \qquad (3)$$

Где Y_n - фактическое значение в последней n-й точке ряда (конечный уровень ряда);

$\hat{Y}_{n+T}$ - прогнозная оценка значения (n + T)-го уровня временного ряда;

$\overline{\Delta Y}$ - значение среднего абсолютного прироста, рассчитанное для временного ряда y_1, y_2 , y_3…. $Y_т$.

T - период упреждения.

Экстраполяция по среднему абсолютному приросту может быть выполнена в том случае, если считать общую тенденцию развития явления линейной, либо выполнить следующее неравенство:

$$\sigma^2_{ост} \le p^2, \quad где \quad p^2 = \frac{1}{2} \cdot \frac{\sum \Delta_i}{n}$$

где $\sigma^2_{ост}$ - остаточная дисперсия, не объясненная эстраполяцией по среднему абсолютному приросту; $\sum \Delta_i$ - общий прирост показателя от начального уровня до конечного y_i.

Продемонстрируем использование рассмотренного приема на следующем примере.

Пример:

По данным, приведенным в таблице требуется обосновать правомерность использования среднего абсолютного прироста для получения прогнозной оценки производства цемента в 2010 году.

а)Рассчитать значение абсолютного прироста и дать его экономическую интерпретацию.

б) определить прогнозное значение производства цемента в 2010г с помощью среднего абсолютного прироста.

Динамика производства цемента млн.тонн					
	2005г	2006г	2007г	2008г	2009г
Объем производства	26	28,5	32,4	35,3	37,7

1. Рассчитаем цепные абсолютные приросты:

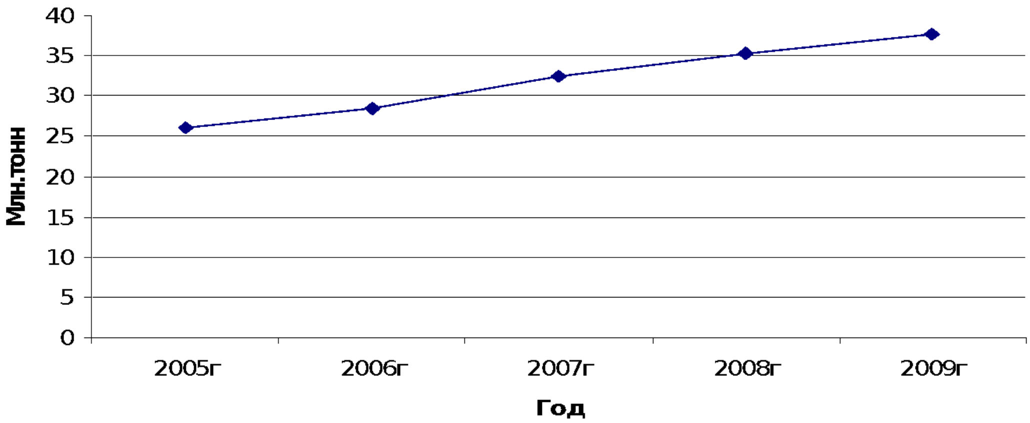
$\Delta y_2 = 28{,}5 - 26{,}0 = 2{,}5$ (млн т); $\Delta y_3 = 32{,}4 - 28{,}5 = 3{,}9$ (млн т);

$\Delta y_4 = 35{,}3 - 32{,}4 = 2{,}9$ (млн т); $\Delta y_5 = 37{,}7 - 35{,}3 = 2{,}4$ (млн т).

В исследуемом периоде цепные абсолютные приросты незначительно изменяются, варьируя от 2,4 до 3,9 млн

Графический анализ свидетельствует о близости процесса к линейному.

Поэтому для определения прогнозного значения показателя в 2010 г. используем средний абсолютный прирост. Немаловажным при выборе этого приема прогнозирования являлось то, что, по мнению экспертов, инерционность развития показателя должна была сохраниться для периода упреждения.



2. Значение среднего абсолютного прироста определим по формуле 2:

$$\overline{\Delta Y} * \frac{37{,}7 - 26{,}0}{4} = 2{,}9$$

млн т ,

т.е., в среднем ежегодно в исследуемом периоде производство цемента увеличивалось на 2,9 (млн т).

3. Определим прогнозное значение производства цемента в 2010г с помощью формулы 3:

$$\hat{Y}_{2010} = Y_{2009} + \overline{\Delta Y} = 37{,}7 + 2{,}9 = 40{,}6 \text{ (млн т)}.$$

$$\hat{Y}_{2014} = Y_{2009} + 5 \cdot \overline{\Delta Y} = 37{,}7 + 5 * 2{,}9 = 37{,}7 + 14{,}5 = 52{,}2 \text{ (млн т)}.$$

Полученный прогноз оказался достаточно точным. В 2010г производство цемента составило 41 млн т, относительная ошибка прогноза - менее 1%. Однако в дальнейшем была нарушена инерционность развития исследуемого показателя, в производстве цемента наблюдался достаточно стремительный рост, линейность развития была нарушена, требовались иные подходы для прогнозирования.

Темп роста характеризует отношение двух сравниваемых уровней ряда, как правило, выраженное в процентах.

Цепной темп роста равен отношению текущего уровня (y_r,) и предыдущего (y_{r-1}):

$$Tt = \frac{y_t}{y_{t-1}} \cdot 100 \, \%, \tag{4}$$

где t = 2, 3, ...n; n - число уровней или длина временного ряда.

Базисный темп роста может быть представлен в виде:

$$T_t = \frac{y_t}{y_b} \cdot 100\%, \tag{5}$$

где y_t, - текущий уровень временного ряда;

y_b - уровень временного ряда, принятый за базу сравнения.

Темп роста всегда положителен. Если темп роста равен 100%, то значение уровня не изменилось, если меньше 100%, то значение уровня понизилось, больше 100%-повысилось.

Средний темп роста является обобщающей характеристикой динамики и отражает интенсивность изменения уровней ряда. Он показывает, сколько в среднем процентов последующий уровень составляет от предыдущего на всем периоде наблюдения. Этот показатель рассчитывается по формуле средней геометрической из цепных темпов роста:

$$\overline{T} = \sqrt[n-1]{T_2 \cdot T_3 \cdot \ldots \cdot T_n} \tag{6}$$

Выразив цепные темпы роста $Г_2$, $Г_3$, ..., T_n через соответствующие уровни ряда, получим:

$$\overline{T} = \sqrt[n-1]{\frac{y_2 \cdot y_3 \cdot \ldots \cdot y_n}{y_1 \cdot y_2 \cdot \ldots \cdot y_{n-1}}} \cdot 100 = \sqrt[n-1]{\frac{y_n}{y_1}} \cdot 100 \, \%. \tag{7}$$

В таблице 1 в графе «Темп роста» приведены выражения для вычисления рассмотренных аналитических показателей динамики (цепных и базисных темпов роста, а также среднего темпа роста).

Прогноз по среднему темпу роста может осуществиться в случае, когда есть основания считать, что общая тенденция ряда динамики характеризуется показательной (экспоненциальной) кривой. Прогнозируемый уровень ряда в этом случае определяется следующей формулой:

$$\hat{y}_{i+T} = y_i \cdot \overline{T}_p^{\,T},$$

где $\overline{T}_p$ - средний темп роста, рассчитанный по формуле средней геометрической.

T- период упреждения.

Доверительный интервал прогноза по среднему темпу роста может быть получен только в том случае, когда средний темп роста определяется с помощью статистического оценивания параметров экспоненциальной кривой.

Пример:

Динамика производства металла млн.тонн					
	2005г	2006г	2007г	2008г	2009г
Объем производства	123	127,5	132,4	141,3	157,7

Рассчитаем средний темп роста:

$$\overline{T} = \sqrt[n-1]{\frac{y_2 \cdot y_3 \cdot \ldots \cdot y_n}{y_1 \cdot y_2 \cdot \ldots \cdot y_{n-1}}} \cdot 100 = \sqrt[n-1]{\frac{y_n}{y_1}} \cdot 100 = \sqrt[5]{\frac{Y_5}{Y_1}} \cdot 100\% = 105{,}09\% \ (1{,}0509)$$

$У_{2010}=У_{2009}\ \overline{T}$ =157,7 *1,0509= **165,7361** млн.тонн или

$У_{2010}= У_{2005}\ \overline{T}^{\,6}$= 123 *1,0509^6= **165,7361** млн.тонн .

Темп прироста характеризует абсолютный прирост в относительных величинах. Определенный в процентах темп прироста показывает, на сколько процентов изменился сравниваемый уровень по отношению к уровню, принятому за базу сравнения. Темп прироста есть выраженное в процентах отношение абсолютного прироста к уровню, принятому за базу сравнения.

Цепной темп прироста может быть представлен в виде:

$$K_t = \frac{y_t - y_{t-1}}{y_{t-1}} \cdot 100 \ \%, \qquad (8)$$

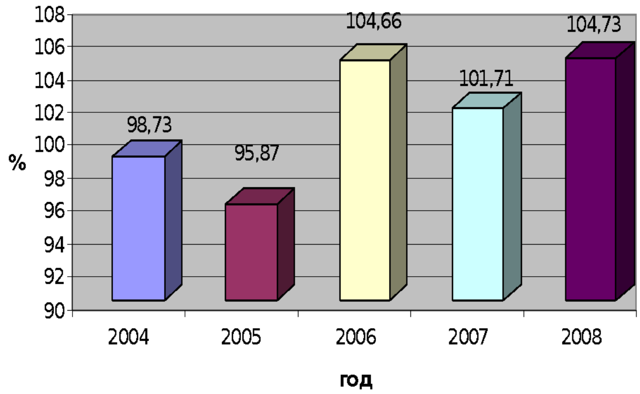
где y_t - текущий уровень временного ряда; t = 2, 3, ... n; n - число уровней или длина временного ряда.

Преобразовав выражение (8), можно показать зависимость цепного темпа прироста от соответствующего темпа роста:

$$K_t = \frac{y_t}{y_{t-1}} \cdot 100 \% - 100\% = T_r - 100\% \qquad (9)$$

где T_r - цепной темп роста.

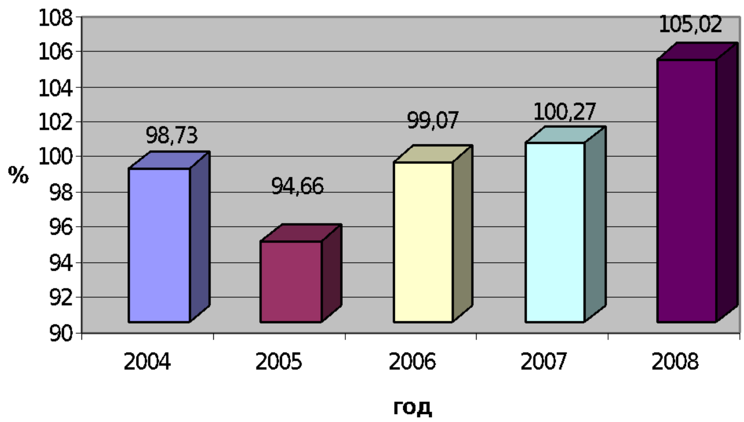
Ценые темпы роста числа малых предприяти РФ



Базисны … солютного прироста к уровню ряда, принятому за базу сравнения:

$$K_r^b = \frac{\Delta y_r^b}{y_b} \qquad (10)$$

Базисные темпы роста числа малых предприяти РФ



Задание к лабораторной работе:

Имеется временные ряды значений изучаемого показателя (задание выполняется в соответствии со своим вариантом из Приложения 1.

1. Построить график исходных данных.

2. Определить основные показатели динамики исходных данных.
Построить графики абсолютных и относительных приростов, абсолютных и относительных темпов роста. . В этой же системе координат построить прогнозы на основании показателей динамики.

3. Сделать вывод о том, какой показатель наиболее точно отражает динамику изменения исходных данный, сделать выводу о характере и динамике развития процесса.

Пример:

Прогноз динамики ряда с помощью среднего абсолютного прироста

Прогнозирование по данным, приведенным в Таблице 1 . Требуется обосновать правомерность использования среднего абсолютного прироста для получения прогнозной оценки производства промышленной продукции. Рассчитаем абсолютные приросты (Таблица 1).

В исследуемом периоде цепные абсолютные приросты незначительно изменяются, варьируя от -45 до 130 млн

Графический анализ свидетельствует о близости процесса к линейному.

Поэтому для определения прогнозного значения показателя в 2010 г. используем средний абсолютный прирост. Немаловажным при выборе этого приема прогнозирования являлось то, что, по мнению экспертов, инерционность развития показателя должна была сохраниться для периода упреждения.

Таблица 1 – Исходные данные о производстве промышленной продукции

Период	Пр-во пром. Прод.(У)	Абсолютный прирост	
		Цепной	Базисный
Январь 2007	780,0		
Февраль 2007	870,0	90,0	90,0
Март 2007	830,0	-40,0	50,0
Апрель 2007	950,0	120,0	170,0
Май 2007	910,0	-40,0	130,0
Июнь 2007	865,0	-45,0	85,0
Июль 2007	900,0	35,0	120,0
Август 2007	940,0	40,0	160,0
Сентябрь 2007	960,0	20,0	180,0
Октябрь 2007	920,0	-40,0	140,0
Ноябрь 2007	1000,0	80,0	220,0
Декабрь 2007	990,0	-10,0	210,0
Январь 2008	1020,0	30,0	240,0
Февраль 2008	1150,0	130,0	370,0
Март 2008	1120,0	-30,0	340,0
Апрель 2008	1230,0	110,0	450,0
Май 2008	1245,0	15,0	465,0
Июнь 2008	1250,0	5,0	470,0

2. Значение среднего абсолютного прироста определим:

$$\overline{\Delta Y} = \frac{1250 - 780}{17} = 27{,}65$$ млн т ,

т.е., в среднем ежемесячно в исследуемом периоде производство промышленной продукции увеличивалось на 27,65 (млн т).

3. Определим прогнозное значение производства промышленной продукции в

$\hat{Y}_{июль\_2008} = Y_{июнь\_2008} + \overline{\Delta Y} =$ 1250 + 27,65= 1277,65 (млн т).

$\hat{Y}_{июль\_2008} = Y_{июнь\_2008} + 2\overline{\Delta Y} =$ 1250+ 2*27,65= 1305,3 (млн т).

$\hat{Y}_{июль\_2008} = Y_{июнь\_2008} + 3\overline{\Delta Y} =$ 1250+ 3*27,65= 1332,95 (млн т).

Прогноз по среднему темпу роста может осуществиться в случае, когда есть основания считать, что общая тенденция ряда динамики характеризуется показательной (экспоненциальной) кривой. Прогнозируемый уровень ряда в этом случае определяется следующей формулой:

$$\hat{y}_{i+T} = y_i \cdot \overline{T}_p^{\,T},$$

где $\overline{T}_p$ - средний темп роста по формуле средней геометрической.

Т- период упреждения.

Доверительный интервал прогноза по среднему темпу роста может быть получен только в том случае, когда средний темп роста определяется с помощью статистического оценивания параметров экспоненциальной кривой.

Рассчитаем средний темп роста:

$$\overline{T} = \sqrt[n-1]{\frac{y_2 \cdot y_3 \cdot \ldots \cdot y_n}{y_1 \cdot y_2 \cdot \ldots \cdot y_{n-1}}} \cdot 100 = \sqrt[n-1]{\frac{y_n}{y_1}} \cdot 100 = \sqrt[17]{\frac{Y_{18}}{Y_1}} \cdot 100\% = 102{,}81\% \ (2{,}8129)$$

$$\hat{Y}_{июль\_2008} = Y_{июнь\_2008} * \overline{T} = 1250*2{,}8129 = 1285{,}13 \text{ (млн т).}$$

$$\hat{Y}_{июль\_2008} = Y_{июнь\_2008} * \overline{T}^2 = 1250*2{,}8129^2 = 1321{,}31 \text{ (млн т).}$$

$$\hat{Y}_{июль\_2008} = Y_{июнь\_2008} * \overline{T}^3 = 1250*2{,}8129^3 = 1358{,}48 \text{ (млн т).}$$

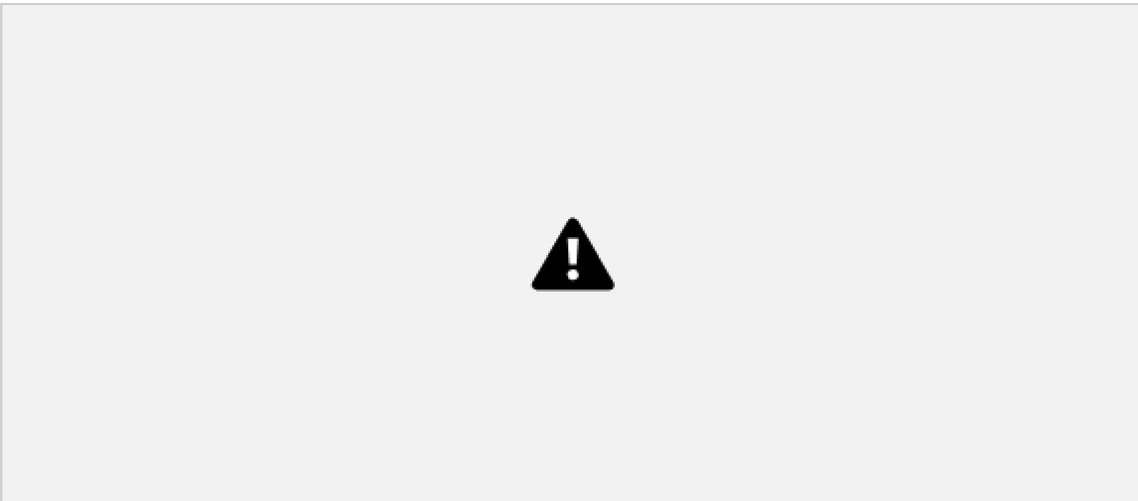


Рисунок 1 – Прогнозы по среднему темпу роста, абсолютному приросту и исходные данные.

Темп прироста характеризует абсолютный прирост в относительных величинах. Определенный в процентах темп прироста показывает, на сколько процентов изменился сравниваемый уровень по отношению к уровню, принятому за базу сравнения. Темп прироста есть выраженное в процентах отношение абсолютного прироста к уровню, принятому за базу сравнения.

Таблица 2 - Расчет показателей динамики

Период	Пр-во пром. Прод.(У)	Темп роста		Темп прироста	
		Цепной	Базисный	Цепной	Базисный
Январь 2007	780,0				
Февраль 2007	870,0	111,5385	111,53	11,53846154	11,5384615
Март 2007	830,0	95,4023	106,41	-4,597701149	6,41025641
Апрель 2007	950,0	114,4578	121,79	14,45783133	21,7948718
Май 2007	910,0	95,78947	116,66	-4,210526316	16,6666667
Июнь 2007	865,0	95,05495	110,89	-4,945054945	10,8974359
Июль 2007	900,0	104,0462	115,38	4,046242775	15,3846154
Август 2007	940,0	104,4444	120,51	4,444444444	20,5128205
Сентябрь 2007	960,0	102,1277	123,07	2,127659574	23,0769231
Октябрь 2007	920,0	95,83333	117,94	-4,166666667	17,9487179
Ноябрь 2007	1000,0	108,6957	128,205	8,695652174	28,2051282
Декабрь 2007	990,0	99	126,92	-1	26,9230769
Январь 2008	1020,0	103,0303	130,76	3,03030303	30,7692308
Февраль 2008	1150,0	112,7451	147,43	12,74509804	47,4358974
Март 2008	1120,0	97,3913	143,58	-2,608695652	43,5897436
Апрель 2008	1230,0	109,8214	157,692	9,821428571	57,6923077
Май 2008	1245,0	101,2195	159,61	1,219512195	59,6153846
Июнь 2008	1250,0	100,4016	160,25	0,401606426	60,2564103

Сглаживание данных

Произведем сглаживание для исходных данных методами: использование скользящей средней, взвешенной скользящей средней и экспоненциальным сглаживанием.

Метод простой скользящей средней применим, если графическое изображение динамического ряда напоминает прямую. Если для о процесса характерно нелинейное развитие, то простая скользящая средняя может привести к существенным искажениям. Когда тренд выравниваемого ряда имеет изгибы и для исследователя желательно сохранить мелкие волны, то целесообразно использовать взвешенную скользящую среднюю.

Сглаживание методом скользящей средней происходит при l=5:

$$\tilde{y}_3 = (y_1 + y_2 + y_3 + y_4 + y_5)/5$$

$$\tilde{y}_4 = (y_2 + y_3 + y_4 + y_5 + y_6)/5$$ и тд.

Восстановление потерянных значений произведем с помощью простой скользящей средней. Для этого вычислим средний абсолютный прирост на последнем и на первом активных участках;

$$\Delta\overline{y}_{\textit{последний}} = (y_{18} - y_{14})/4 = 25;$$
$$\Delta\overline{y}_{\textit{первый}} = (y_5 - y_1)/4 = 32{,}5.$$

Последовательно прибавим средние абсолютные приросты к последнему и первому сглаженному значению соответственно.

Сглаживание методом взвешенной скользящей средней при l=5 происходит:

$$\tilde{y}_3 = \frac{1}{35}(-3y_1 + 12y_2 + 17y_3 + 12y_4 - 3y_5)$$

$$\tilde{y}_4 = \frac{1}{35}(-3y_2 + 12y_3 + 17y_4 + 12y_5 - 3y_6)$$ и тд.

Восстановление потерянных значений аналогично как и в предыдущем сглаживании.

Экспоненциальное сглаживание, как и скользящее среднее, используется для выравнивания (сглаживания) значений временных рядов. Если имеются дискретные наблюдения y_1, y_2, ..., y_n, то сглаженные значения вычисляются по формуле:

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t \, .$$

где y_t — сглаженное значение для предыдущего t, α - постоянная сглаживания, также называемая *фактором затухания* (это число из интервала (0, 1)).

Экспоненциальное сглаживание при α =0,3:

$$\overline{y} = (\sum_{i=1}^{18} y_i)/18 = 996{,}1;$$

$$\hat{y}_1 = y_1 * \alpha + (1-\alpha)\overline{y} = 931{,}3.$$

$$\hat{y}_2 = y_2 * \alpha + (1-\alpha)\hat{y}_1 = 912{,}9.$$ и тд.

Построим прогноз среднего темпа роста по сглаженным значениям метода экспоненциального сглаживания.

$$\overline{T} = \sqrt[17]{\frac{y_{18}}{y_1}} \cdot 100$$ =101,4(1,014)

$$\hat{y}_{19} = \hat{y}_{18}\overline{T}^1 = 1196{,}2$$

$$\hat{y}_{20} = \hat{y}_{18}\overline{T}^2 = 1221{,}9$$

$$\hat{y}_{21} = \hat{y}_{18}\overline{T}^3 = 1229{,}9$$

Все расчеты приведены в таблице 3.

Таблица 3 – Расчет значений сглаживания скользящей средней

Период	Пр-во пром. Прод.(У)	Скользящее среднее l=5	Экспоненциальное Y1*	Взвешенные скольз. средние
Январь 2007	780,0	803,00	931,3	817,3
Февраль 2007	870,0	835,50	912,9	849,8
Март 2007	830,0	868,00	888,0	882,29
Апрель 2007	950,0	885,00	906,6	909,29
Май 2007	910,0	891,00	907,6	916,00
Июнь 2007	865,0	913,00	894,8	878,71
Июль 2007	900,0	915,00	896,4	895,71

Продолжение таблицы 3

Период	Пр-во пром. Прод.(У)	Скользящее среднее l=5	Экспоненциальное Y1*	Взвешенные скольз. средние
Август 2007	940,0	917,00	909,5	941,29
Сентябрь 2007	960,0	944,00	924,6	941,14

Октябрь 2007	920,0	962,00	923,2	953,43
Ноябрь 2007	1000,0	978,00	946,3	970,86
Декабрь 2007	990,0	1016,00	959,4	996,00
Январь 2008	1020,0	1056,00	977,6	1047,43
Февраль 2008	1150,0	1102,00	1029,3	1102,00
Март 2008	1120,0	1153,00	1056,5	1165,86
Апрель 2008	1230,0	1199,00	1108,6	1202,57
Май 2008	1245,0	1224,00	1149,5	1227,6
Июнь 2008	1250,0	1249,00	1179,6	1252,6
Июль 2008		1274,00	1196,2	1277,6
Август 2008		1299,00	1212,9	1302,6
Сентябрь 2008		1324,00	1229,9	1327,6

Построим графики полученных прогнозов:

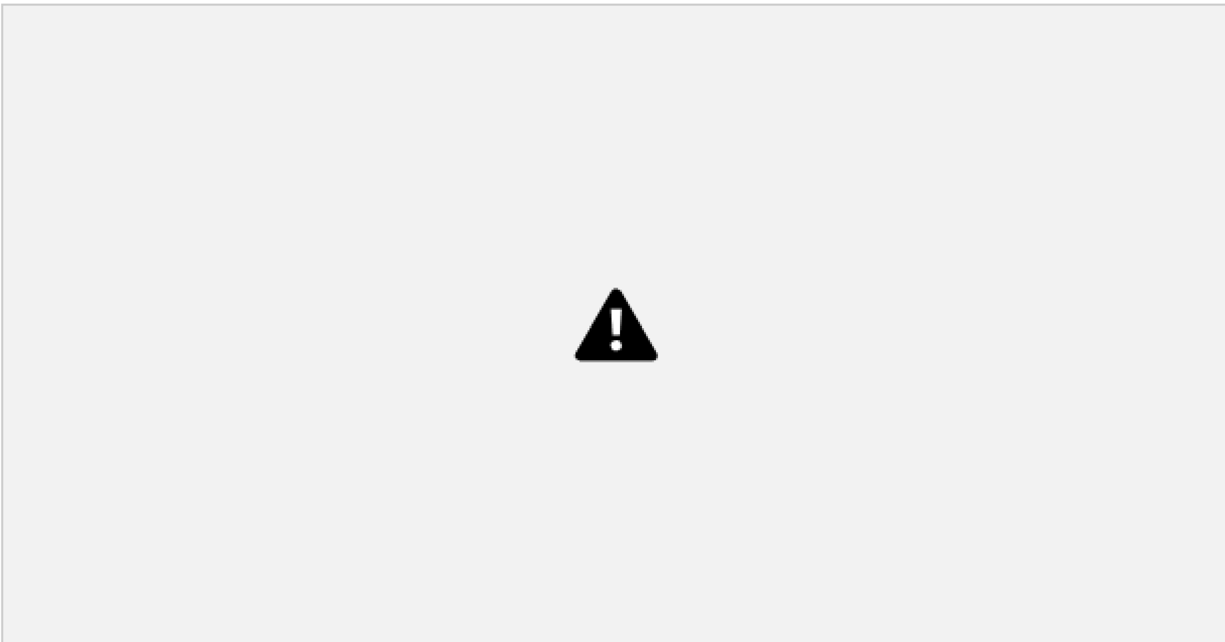


Рисунок 2 - Сглаживание простой скользящей средней L=5

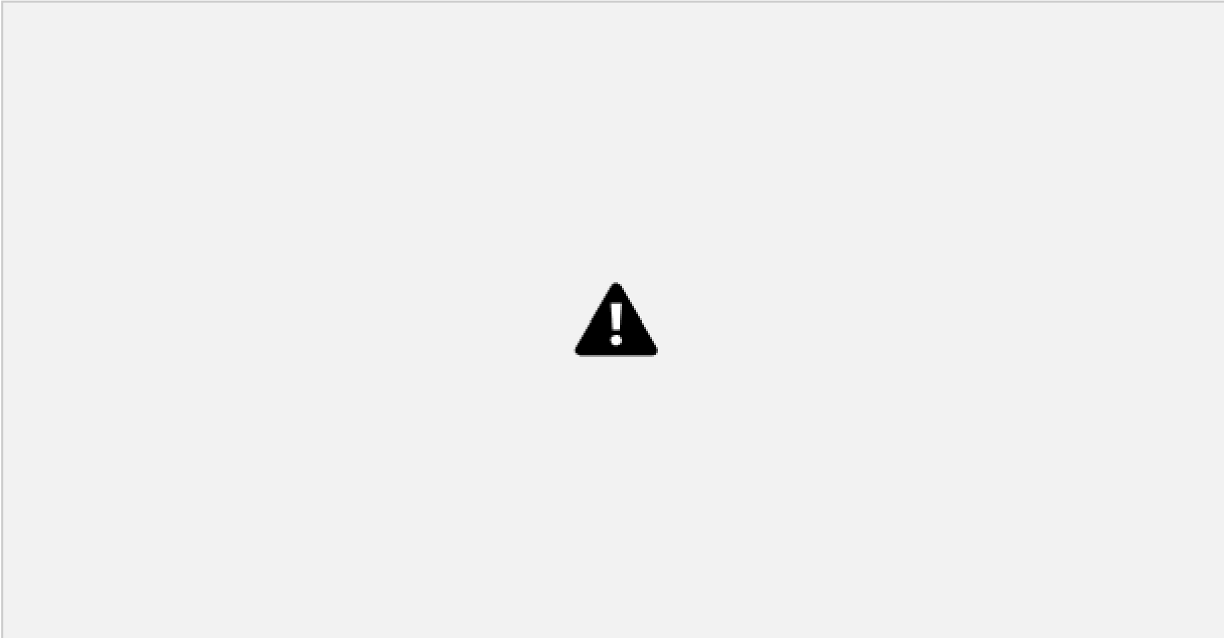


Рисунок 3 - Сглаживание взвешенной скользящей средней L=5

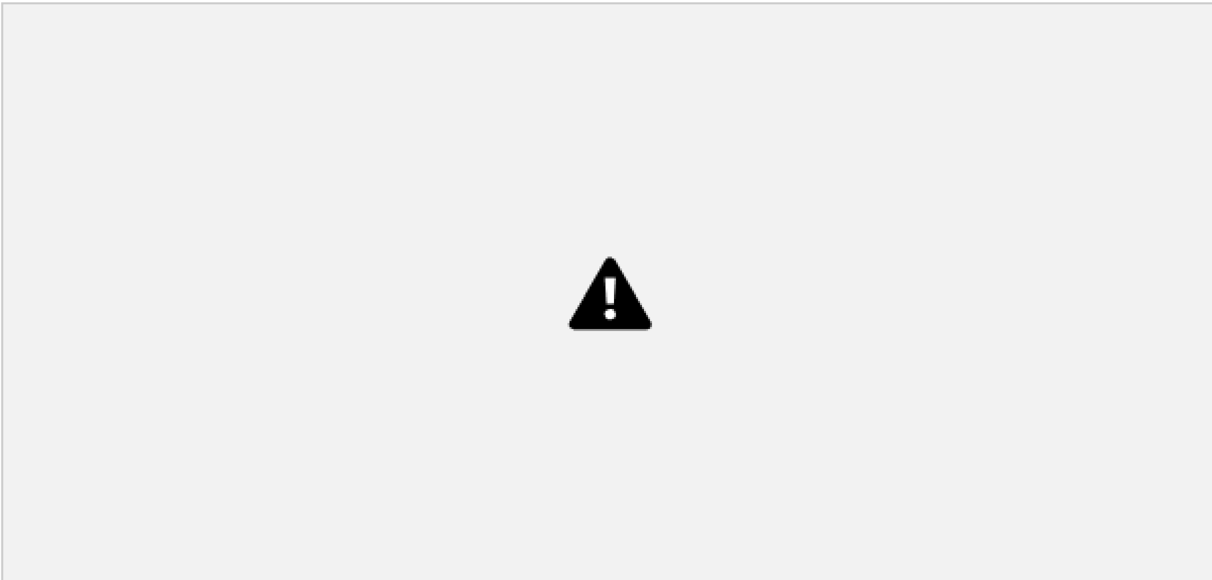


Рисунок 4 – Экспоненциальное сглаживание

В связи стем, что данные имеют линейный характер, прогноз полученный по экспоненциальному сглаживанию не достоверен, что и видно на рисунке 4. А оставшиеся прогнозы отличабтся незначительно и достаточно хорошо отражают связь прогноза с исходнями данными. Получив прогнозируемое значение производства промышленных товаров на сентябрь 2008 1324 и 1327,6 .

Лабораторная работа № 2

Прогнозирование тенденции развития с помощью моделей кривых роста.

Цель работы: ознакомиться с методами сглаживания исходных данных кривыми роста. Научиться проводить проверку адекватности выбранной модели исследуемому процессу и выполнять оценку точности модели. Стоить точечный и интервальный прогноз на основе полученной модели.

Общие положения

1. Сглаживание кривыми роста.

Основная цель создания трендовых моделей экономической динамики — на их основе сделать прогноз о развитии изучаемого процесса на предстоящий промежуток времени. Прогнозирование на основе временного ряда экономических показателей относится к одномерным методам прогнозирования, базирующимся на экстраполяции, т.е. на продлении на будущее тенденции, наблюдавшейся в прошлом. Рассмотрим метод экстраполяции на основе так называемых кривых роста экономической динамики.

Кривые роста - это математические функции, предназначенные для аналитического выравнивания временного ряда.

Наиболее часто в практической работе используются кривые роста, которые позволяют описывать процессы трех основных типов: без предела роста; с пределом роста без точки перегиба; с пределом роста и точкой перегиба.

Таблица1 - Виды кривых роста

№ п/п	Аналитическое выражение функции	Описание
1	$\hat{y}_t = \alpha_0 + \alpha_1 \cdot t$	- прямая
2	$\hat{y}_t = \alpha_0 + \alpha_1 \cdot t + \alpha_2 \cdot t^2$	- парабола II порядка
3	$\hat{y}_t = \alpha_0 + \alpha_1 \cdot t + \alpha_2 \cdot t^2 + \alpha_3 \cdot t^3$	- парабола III порядка

4	$\hat{Y}_t = a \cdot b^t \quad \hat{Y}_t = c + a \cdot b^t$	- простая и модифицированная экспонента
5	$\hat{y}_t = a \cdot b^t$	-показательная кривая
6	$\hat{y}_t = a + b / t$	-гипербола
7	$\hat{y}_t = b \cdot c^{a^t}$	-кривая Гомперца
8	$\hat{y}_t = \dfrac{k}{c + b \cdot a^{-t}}$	-логистическая кривая кривая Перла-Рида

Наиболее распространенными моделями кривых роста являются полиномы первого и второго порядков.

Формулы для определения коэффициентов линейной и параболической моделей можно представить следующим образом:

Таблица2 - Определение коэффициентов кривых роста

№ п/п	Вид модели	Формулы для определения коэффициентов
1	Линейная модель $\hat{y}_t = a_0 + a_t \cdot t$	$a_0 = \dfrac{\sum y_t}{n} a_1 = \dfrac{\sum y_t \cdot t}{\sum t^2}$
2	Параболическая модель $\hat{y}_t = a_0 + a_t \cdot t + a_2 \cdot t^2$	$a_0 = \dfrac{\sum y_t}{n} - \dfrac{\sum t^2}{n}\left[\dfrac{n \cdot \sum y_t t^2 - \sum t^2 \cdot \sum y_t}{n \cdot \sum t^4 - (\sum t^2)}\right]$ $a_1 = \dfrac{\sum y_t \cdot t}{\sum t^2}$ $a_2 = \dfrac{n \cdot \sum y_t t^2 - \sum t^2 \cdot \sum y_t}{n \cdot \sum t^4 - (\sum t^2)}$

Для расчета коэффициентов модели тренда третьего порядка воспользуемся следующей формулой:

$A = (Z^т \cdot Z)^{-1} \cdot Z^т \cdot Y$, где

$$A=\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}; \; Z=\begin{pmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 1 & t_{19} & t_{19}^2 & t_{19}^3 \end{pmatrix}\begin{pmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 1 & t_{19} & t_{19}^2 & t_{19}^3 \end{pmatrix}; \; Z^T –$$

транспонированная матрица Z; $Y=Y_t$.

При моделировании экономического процесса, заданного временным рядом, путем сглаживания исходного ряда с помощью кривых роста необходима проверка соответствия выбранной модели реальному процессу (явлению).

Процесс моделирования включает следующие этапы:

-предварительный анализ данных;

- определения наличия тренда;

- в случае наличия тренда отбор одной или нескольких кривых роста и определения их параметров для получения трендовых моделей исходного временного ряда;

- проверка адекватности модели анализируемым данным, обоснование использования этих моделей для анализа и прогнозирования изучаемого экономического процесса (явления.);

- определение точности модели.

2. Пример выполнения задания

В таблице представлены данные о прибыли компании по кварталам. Требуется рассчитать прогноз прибыли в следующем квартале, предположив. Что тенденция ряда может быть описана:

I)параболической моделью

II)параболической моделью третьего порядка

I) Для расчета коэффициентов параболического тренда воспользуемся выражениями в таблице, полученными из системы уравнений после переноса в начала координат. Так как число уровней нечетное n=19, то центральный уровень (10) принимается за начало отсчета, ему соответствует t=0. Вышестоящие уровни нумеруем с шагом -1, нижестоящие - с шагом +1. Данные представлены в таблице3.

$$a_0 = 37{,}17895 - \frac{570}{19} \cdot -0{,}23869 = 44{,}33972 \qquad a_1 = \frac{-2036{,}9}{570} = -3{,}57351$$

$$a_2 = \frac{19 \cdot 17953{,}9 - 570 \cdot 706{,}4}{19 \cdot 30666 - (570)^2} = \text{-}0{,}23869$$

Следовательно, уравнение параболического тренда имеет вид:

$$\hat{y}_t = 44{,}33972 \text{ - } 3{,}57351 \cdot t \text{ - } 0{,}23869 \cdot t^2$$

Таблица 3 –Расчеты для определения коэффициентов модели

№ п/п	Уровень инфляции (y1)	t	Yt*t	t2	Yt*t2	t4
1	22,1	-9	-198,9	81	1790,1	6561
2	38,9	-8	-311,2	64	2489,6	4096
3	32,5	-7	-227,5	49	1592,5	2401
4	98,1	-6	-588,6	36	3531,6	1296
5	130	-5	-650	25	3250	625
6	87,4	-4	-349,6	16	1398,4	256
7	13,8	-3	-41,4	9	124,2	81
8	84	-2	-168	4	336	16
9	54,2	-1	-54,2	1	54,2	1
10	22,7	0	0	0	0	0
11	20,2	1	20,2	1	20,2	1
12	18,6	2	37,2	4	74,4	16
13	15,1	3	45,3	9	135,9	81
14	12	4	48	16	192	256
15	11,7	5	58,5	25	292,5	625
16	10,9	6	65,4	36	392,4	1296
17	9	7	63	49	441	2401
18	11,9	8	95,2	64	761,6	4096
19	13,3	9	119,7	81	1077,3	6561
Сумма	706,4		-2036,9	570	17953,9	30666

Согласно этой модели оценка уровня ряда при t=0 равна 44,33972. Это расчетное значение превышает фактическое, равное 22,7.

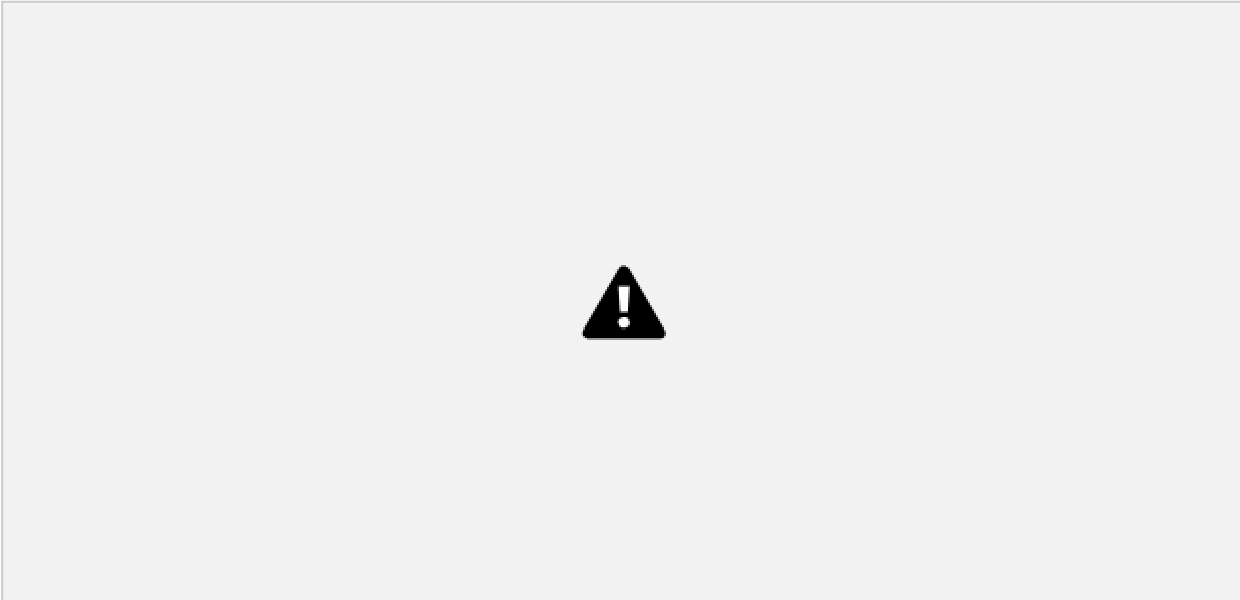
В соответствии с моделью ежеквартальный прирост прибыли компании в среднем составляет около 7,5 тыс.$.

Для прогнозирования на базе полученной модели необходимо подставить в нее соответствующее значение времени t=10 (при переносе начала отсчета), t=20 (без переноса).

Рассчитанное прогнозное значение:

$$\hat{y}_{10} = 44{,}33972 - 3{,}57351 \cdot 10 - 0{,}23869 \cdot 10^2 = -15{,}2646$$

Графическая иллюстрация:



II) Для расчета коэффициентов модели тренда третьего порядка воспользуемся следующей формулой:

$A = (Z^{т} \cdot Z)^{-1} \cdot Z^{т} \cdot Y$, где

$A = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}$; $Z = \begin{pmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 1 & t_{19} & t_{19}^2 & t_{19}^3 \end{pmatrix}\begin{pmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 1 & t_{19} & t_{19}^2 & t_{19}^3 \end{pmatrix}$; Z^{T} – транспонированная матрица Z; $Y = Y_t$.

$A = \begin{pmatrix} 44{,}33972 \\ -10{,}9895 \\ -0{,}23869 \\ 0{,}137844 \end{pmatrix}\begin{pmatrix} 44{,}33972 \\ -10{,}9895 \\ -0{,}23869 \\ 0{,}137844 \end{pmatrix}$

Следовательно, уравнение модели тренда третьего порядка имеет вид:

$$\hat{y} = 44{,}33972 - 10{,}9895 \cdot t - 0{,}23869 \cdot t^2 + 0{,}137844 \cdot t^3$$

Графическая иллюстрация:

Таблица 4

Параболическая модель	Полиномиальная модель третьего порядка	Исходный ряд
57,16721805	23,42310321	22,1
57,6514787	46,40344042	38,9
57,65835471	62,28989988	32,5
57,18784609	71,90954324	98,1
56,23995282	76,08943214	130
54,81467492	75,6566282	87,4
52,91201238	71,43819308	13,8
50,53196521	64,2611884	84
47,67453339	54,95267581	54,2
44,33971694	44,33971694	22,7
40,52751585	33,24937343	20,2
36,23793012	22,50870693	18,6
31,47095975	12,94477906	15,1
26,22660475	5,384651468	12
20,5048651	0,655385791	11,7
14,30574082	-0,415956335	10,9
7,629231903	2,99768673	9
0,475338346	11,72337662	11,9
-7,15593985	26,58817498	13,3

Рассчитанное прогнозное значение:

$$\hat{y} = 44{,}33972 - 10{,}9895 \cdot 10 - 0{,}23869 \cdot 10^2 + 0{,}137844 \cdot 10^3 =$$

$$\hat{y} = 44{,}33972 - 10{,}9895 \cdot 10 - 0{,}23869 \cdot 10^2 + 0{,}137844 \cdot 10^3 = 48{,}41914$$

Таблица 5

№ п/п	E(t)	St	пики E

1	-35,0672	+	-
2	-18,7515	-	1
3	-25,1584	+	1
4	40,91215	+	0
5	73,76005	-	1
6	32,58533	-	0
7	-39,112	+	1
8	33,46803	-	1
9	6,525467	-	0
10	-21,6397	+	1
11	-20,3275	+	0
12	-17,6379	+	0
13	-16,371	+	0
14	-14,2266	+	0
15	-8,80487	+	0
16	-3,40574	+	0
17	1,370768	+	0
18	11,42466	+	0
19	20,45594		-

Таблица 6

№ п/п	e(t)	пики e	St
1	-1,3231	-	-
2	-7,50344	0	-
3	-29,7899	1	+
4	26,19046	0	+
5	53,91057	1	-
6	11,74337	0	-
7	-57,6382	1	+
8	19,73881	1	-
9	-0,75268	0	-
10	-21,6397	1	+
11	-13,0494	0	+
12	-3,90871	0	+
13	2,155221	0	+
14	6,615349	0	+
15	11,04461	0	+
16	11,31596	1	-
17	6,002313	0	-
18	0,176623	0	-
19	-13,2882	-	

Проверим случайность колебаний остаточных отклонений с помощью критерия серий. Для того, чтобы последовательность E(t) была случайной протяженность самой длинной серии (S_{max}) не должна быть большой, а число серий (υ) малым. Для проверки этого условия используются соотношения:

1) $S_{max} \leq S_0$
2) $\upsilon > [\frac{2(19-2)}{3} - 1{,}96\sqrt{(16*19-29)/90}]$

$\upsilon > [\frac{2(19-2)}{3} - 1{,}96\sqrt{(16*19-29)/90}]$

Если хотя бы одно из неравенств не выполняется, то свойство случайности не выполняется.

I.При n=19, S_{max}=9 и υ=7 получаем:

1) 9>5
2) 7>5

II. При n=19, S_{max}=9 и υ=7 получаем:

1) 6>5
2) 7>5

Одно из условий не выполнено, следовательно, колебания остаточных отклонений не случайны, а модели не корректны.

Проверку случайности уровней ряда остатков проведем на основе критерия поворотных точек (критерия пиков). В соответствии с этим критерием каждый уровень ряда сравнивается с двумя рядом стоящими. Если он больше или меньше их, то эта точка считается поворотной. Далее подсчитываем сумму поворотных точек «P». В случайном ряду чисел должно выполняться неравенство:

$$P > [\frac{2(N-2)}{3} - 2\sqrt{(16N-29)/90}]$$

I. При N=19 и P=6 получим:

$$6 > [\frac{2(19-2)}{3} - 2\sqrt{(16*19-29)/90}]$$

$$6 < 7$$

II. При N=19 и P=6 получим:

$$6 > [\frac{2(19-2)}{3} - 2\sqrt{(16*19-29)/90}]$$

$$6 < 7$$

Следовательно свойство случайности не выполняется для обеих моделей.

Независимость ряда остатков проверим по d-критерию (в качестве критических используем уровни d_1=1,08 и d_2=1,36).

I. $$d = \frac{\sum_{t=2}^{N}[E(t)-E(t-1)]^2}{\sum_{t=2}^{N}E^2(t)} = \frac{19653{,}621}{15385{,}82} = 1{,}277$$

$$d = \frac{\sum_{t=2}^{N}[E(t)-E(t-1)]^2}{\sum_{t=2}^{N}E^2(t)} = \frac{19653{,}621}{15385{,}82} = 1{,}277$$

II. $$d = \frac{\sum_{t=2}^{N}[E(t)-E(t-1)]^2}{\sum_{t=2}^{N}E^2(t)} = \frac{18349{,}508}{9552{,}953} = 1{,}920$$

$$d = \frac{\sum_{t=2}^{N}[E(t)-E(t-1)]^2}{\sum_{t=2}^{N}E^2(t)} = \frac{18349{,}508}{9552{,}953} = 1{,}920$$

Вычисленная величина d сравнивается с двумя табличными. Если d находится в интервале от нуля до d_1, то уровни ряда остатков сильно автокоррелированы, а модель неадекватна. Если значение d попадает в интервал от d_2 до 2, то уровни ряда являются независимыми. Если расчетное значение попадает в интервал между критическими уровнями d_1 и d_2, то однозначного вывода сделать нельзя и необходимо применить другие критерии, например, первый коэффициент автокорреляции r(1), который вычисляется по формуле:

$$r(1) = \frac{\sum_{t=2}^{N}E(t)\cdot E(t-1)}{\sum_{t=2}^{N}E^2(t)}$$

I. $r(1) = \frac{4734{,}934}{15385{,}82} = 0{,}308$ $r(1) = \frac{4734{,}934}{15385{,}82} = 0{,}308$

II. $r(1) = \frac{289{,}036}{9552{,}953} = 0{,}030\, r(1) = \frac{289{,}036}{9552{,}953} = 0{,}030$

Если $|r(1)| > r_{таб} > r_{таб}$, то присутствие в ряду существенной автокорреляции подтверждается.

Нормальность распределения остаточной компоненты по R/S-критерию

$$RS=[E_{max}-E_{min}]:S_{E},$$

где E_{max}-максимальный уровень ряда остатков;

E_{min} – минимальный уровень ряда остатков;

S_E – среднее квадратическое отклонение.

Если расчетное значение RS-критерия попадает между критическими значениями с заданным уровнем вероятности, то гипотеза о нормальном распределении ряда остатков принимается. Для N=19 и 5%-ного уровня значимости этот интервал равен ().

I. E_{max}=73,76005, E_{min}=-35,0672

$$S_E = \sqrt{\sum_{t=1}^{N} E^2(t):(N-1)} = \sqrt{\frac{15385{,}82}{18}} =$$

$$S_E = \sqrt{\sum_{t=1}^{N} E^2(t):(N-1)} = \sqrt{\frac{15385{,}82}{18}} = 29{,}23641$$

RS=(73,76005-(-35,0672)):29,23641=3,722319

Расчетное значение попадает в интервал. Следовательно, остаточная компонента распределена нормально.

II. E_{max}=53,911, E_{min}=-57,638

$$S_E = \sqrt{\sum_{t=1}^{N} E^2(t):(N-1)} = \sqrt{\frac{9552{,}953}{18}} =$$

$$S_E = \sqrt{\sum_{t=1}^{N} E^2(t):(N-1)} = \sqrt{\frac{9552{,}953}{18}} = 23{,}037$$

RS=(53,911-(-57,638)):23,037=4,842

Расчетное значение попадает в интервал. Следовательно, остаточная компонента распределена нормально.

Для проверки нормальности распределения остаточной компоненты используем след соотношения:

$|A|<1,5\sigma_e$

$|Э+6/(n+1)|<1,5\ \sigma_e$

А,Э – асимметрия и эксцесс ряда остатков, которые рассчитываются по формулам:

$$A = \frac{\frac{1}{n}\sum_{t=1}^{n} e_t^3}{\sqrt{(\frac{1}{n}\sum_{t=1}^{n} e_t^2)^3}} - 3$$

$$Э = \frac{\frac{1}{n}\sum_{t=1}^{n} e_t^4}{(\frac{1}{n}\sum_{t=1}^{n} e_t^2)^3} - 3$$

I. Для рассматриваемой модели А= -2,098479; Э= -2,996039; σ_e= 29,236414. Можно сделать вывод, что остатки нормально распределены.

II. А= -3,21398; Э= -2,99121; σ_e= 23,03735. Можно сделать вывод, что остатки нормально распределены.

Проверим равенство нулю математического ожидания ряда остатков, которое осуществляется следующим образом:

1)Выдвигается гипотеза H_0: $\overline{|e|} = 0\ \overline{|e|} = 0$;

2)Рассчитывается статистика t- Стьюдента

$t_p = \frac{\overline{|e|}}{\sigma_e} \cdot \sqrt{n}\ \frac{\overline{|e|}}{\sigma_e} \cdot \sqrt{n}$

Если $t_p<t_{табл.}$, взятого для α=0,05 и числа степеней свободы n-1, то гипотеза о равенстве нулю принимается, в противном случае – отвергается.

I. t_p=0, $t_{табл}$=2,101

$t_p<t_{табл}$

Математическое ожидание равно нулю.

$$\text{II. } t_p=0,\ t_{табл}=2{,}101$$

$$t_p<t_{табл}$$

Математическое ожидание равно нулю.

Рассчитаем ошибки и доверительный интервал прогноза для уровня значимости 5 % (α = 0,05).

Ошибка прогноза – величина, характеризующая расхождение фактического и прогнозного показателя

$$\Delta t = \hat{y}_t - y_t$$

$\hat{y}_t$ - прогнозное значение

Средняя абсолютная ошибка прогноза (MAD)

$$\bar{\Delta t} = \frac{\sum_{i=1}^{n}|\Delta t|}{n}$$

$$\text{I. } \overline{\Delta t} = \frac{441.0048}{19} = 23{,}211\ \overline{\Delta t} = \frac{441.0048}{19} = 23{,}211$$

$$\text{II. } \overline{\Delta t} = \frac{297.7866}{19} = 15{,}673\ \overline{\Delta t} = \frac{297.7866}{19} = 15{,}673$$

На практике используются относительные ошибки, выраженные в процентах

$$\delta_t = \frac{\hat{y}_t - y_t}{y_t} * 100$$

Средняя абсолютная процентная ошибка (MAPE)

$$MAPE = \frac{\sum_{t=1}^{n}|\delta_t|}{n}$$

$$\text{I. } MAPE = \frac{739{,}32}{19} = 38{,}91\ MAPE = \frac{739{,}32}{19} = 38{,}91$$

MAPE<50%. Это свидетельствует об удовлетворительной точности модели.

II. $MAPE = \frac{375,96}{19} = 19,79$ $MAPE = \frac{375,96}{19} = 19,79$

MAPE<20%.это свидетельствует о хорошей точности модели.

Для проведения сравнительной оценки нескольких моделей используется средняя процентная ошибка (MPE)

$$MPE = \frac{1}{n}\sqrt{\frac{\sum_{t=1}^{n}|\hat{y}_t - y_t|}{y_t}} * 100$$

I. $MPE = \frac{1}{19} \cdot \sqrt{\frac{441,0048}{706,4}} * 100\% = 4,16\%$

$MPE = \frac{1}{19} \cdot \sqrt{\frac{441,0048}{706,4}} * 100\% = 4,16\%$

II. $MPE = \frac{1}{19} \cdot \sqrt{\frac{297,7866}{706,4}} * 100\% = 3,42\%$

$MPE = \frac{1}{19} \cdot \sqrt{\frac{297,7866}{706,4}} * 100\% = 3,42\%$

Согласно средней процентной ошибке, вторая модель лучше первой.

Иногда для оценки моделей используются следующие два показателя – сумма квадратов ошибок (SSE) и усреднённое значение квадратов ошибок (MSE)

$$SSE = \sum_{t=1}^{n} e_t^2 \quad \text{и } MSE = \frac{1}{n}\sum_{t=1}^{n} e_t^2$$

I. SSE= 15385,82; MSE= 809,7801

II. SSE= 9552,953; MSE= 502,787

Для расчёта доверительного интервала используют выражение:

$$U(l) = \hat{y} \pm t_\alpha \cdot \hat{\delta}_y \cdot K$$

где $\hat{\delta}_y = \sqrt{\frac{1}{n-k-1}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}$ - СКО тренда n – количество уровней ряда

k – порядок модели (k=1)

y_t - уровни ряда

$\hat{y}_t$ - выровненные уровни, принадлежащие тренду

t_α - табличное значение t-критерия Стьюдента (t_α =2,1199 при α = 0,05)

K – коэффициент, зависящий от длины временного ряда и периода упреждения l =1

$$K = \sqrt{1 + \frac{1}{n} + n\frac{(n + l - \bar{t})^2}{\sum_{t=1}^{n}(t_t - \bar{t})^2}}$$

I. $\widehat{\delta_y} = \sqrt{\frac{1}{19-1-1} * 15385{,}82} = \widehat{\delta_y} = \sqrt{\frac{1}{19-1-1} * 15385{,}82} =$ 30,084;

$$K = \sqrt{1 + \frac{1}{19} + 19 \cdot \frac{(19 + 1 - 0)^2}{570}} = 3{,}793$$

Доверительный интервал прогноза для уровня значимости 5%:

$U(1) = \hat{y} \pm 2{,}101 \cdot 30{,}084 \cdot 3{,}793 = \hat{y} \pm 239{,}735$

II. $\widehat{\delta_y} = \sqrt{\frac{1}{19-1-1} * 9552{,}953} = 23{,}705;$

$\widehat{\delta_y} = \sqrt{\frac{1}{19-1-1} * 9552{,}953} = 23{,}705;$

$$K = \sqrt{1 + \frac{1}{19} + 19 \cdot \frac{(19 + 1 - 0)^2}{570}} = 3{,}793$$

Доверительный интервал прогноза для уровня значимости 5%:

$U(1) = \hat{y} \pm 2{,}101 \cdot 23{,}705 \cdot 3{,}793 = \hat{y} \pm 188{,}904$

Контрольные вопросы

1. Объясните назначение кривых роста. Влияние каких компонент временного ряда устраняется с их помощью?
2. Какие задачи преследует предварительный анализ данных? Какие подходы имеются?
3. Поясните, какие кривые роста целесообразно использовать и для каких временных рядов.
4. Приведите алгоритм сглаживания кривыми роста.
5. Напишите формулы для расчета коэффициентов линейной и параболических моделей кривых роста..
6. Каким образом оценить адекватность полученной модели анализируем данным?
7. Какие имеются показатели для оценки точности модели?
8. Что такое точечное и интервальное прогнозирование?
9. Как получить прогнозное значение на основе кривой роста?

Лабораторная работа № 3

Статистический анализ и прогнозирование на основе тренд-сезонных моделей

Цель работы: ознакомиться с методами оценивания сезонных факторов, сезонной декомпозицией данных и научиться проводить анализ полученных данных и прогнозирование с учетом полученных оценок сезонности в программе Excel.

Общие положения

Процедуры расчета сезонной составляющей зависит от принятой модели временного ряда, содержащей сезонность:

А) в аддитивной модели характеристики сезонности будут складываться со значением тренда и измеряться в абсолютных величинах.

Б) в мультипликативной модели характеристики сезонности будут умножаться на значение тренда и измеряться в относительных величинах.

Схему оценивания сезонной компоненты можно представить в виде следующих последовательных этапов.

I этап. Сглаживание исходного временного ряда с помощью процедуры скользящей средней при четной длине интервала сглаживания для предварительного оценивания тенденции развития.

II этап. Вычисление уровней временного ряда x_t, представляющих собой отношения/разности фактических уровней y_t и сглаженных значений y_t', полученных на предыдущем шаге (соответственно для случая мультипликативной/ аддитивной сезонности).

III этап. Усреднение значений уровней x_t для одноименных месяцев (кварталов) с целью элиминирования влияния случайной составляющей и определения предварительных значений сезонной компоненты.

IV этап. Проведение корректировки первоначальных значений сезонной составляющей с учетом мультипликативного или аддитивного характера сезонности для того чтобы суммарное воздействие сезонности на динамику было нейтральным.

1. Прогнозирование на основе тренд-сезонных моделей

1.Для описания тенденции воспользуемся процедурой скользящей средней при четной длине интервала сглаживания $l=2p$.

Для рядов месячной динамики скользящая средняя при l=12 на каждом активном участке будет определяться выражением:

$$\hat{y} = \frac{\frac{1}{2}y_{t-6} + y_{t-5} + \ldots + y_t + \ldots + y_{t*5} + \frac{1}{2}y_{t*6}}{12}$$

Для рядов квартальной динамики при *l=4.*можно использовать выражение :

$$\hat{y} = \frac{\frac{1}{2}y_{t-2} + y_{t-1} + y_t + y_{t+1} + \frac{1}{2}y_{t+2}}{4}$$

2. Рассчитаем отклонение фактических значений y_t от уровней сглаженного ряда $y_t^{'}$, полученных на предыдущем шаге:

А) в случае аддитивной сезонной составляющей :

$$x_t = y_t - y_t'$$

Б) в случае мультипликативной сезонной составляющей:

$$x_t = \frac{y_t}{y_t'}$$

Уровни вновь полученного ряда x_t отражают эффект сезонности и случайности.

3. Для устранения (элиминирования) влияния случайных факторов определяем предварительные значения сезонной составляющей как средние значения из уровней x_t для одноименных месяцев (кварталов).

4. Проведем корректировку первоначальных значений сезонной составляющей, вызванную тем, что суммарное воздействие сезонности на динамику предполагается нейтральным.

Для *сезонных колебаний мультипликативной форме* выражается в том, что средняя арифметическая из значений коэффициентов сезонности для полного сезонного цикла должна быть равна 1.

$$S_i = \bar{x}_t \cdot k \quad \bar{S}_i = \bar{x}_i k$$ (i=1,2,…,m), где $$k = \frac{m}{\sum_{i=1}^{m} \bar{x}_t}$$

;

m - число фаз в полном сезонном цикле (как правило, m=12 для рядов месячной динамики и m=4 для квартальных данных).

Для *сезонных колебаний в аддитивной форме* взаимопогашаемость сезонных колебаний выражается в том, что для полного сезонного цикла должна быть равна нулю.

Поэтому окончательные оценки коэффициентов сезонности получим с помощью следующего выражения:

$$S_t = \breve{x}_t - \breve{x} \ (i = 1,2,...m) \qquad \text{где} \qquad \overline{x} = \frac{1}{m}\sum_{i=1}^{m} \breve{x}_i ,$$

m – число фаз в полном сезонном цикле.

Пример :

В табл. 1 представлены ежемесячные данные об объемах продаж розничных магазинах США в 2011 – 2014 гг.

Требуется:

1) Провести исследование компонентного состава анализируемого временного ряда;
2) Рассчитать прогнозную оценку суммарного обьема продаж во всех розничных магазинах США в 2015 г.

Графический анализ исходного временного ряда свидетельствует о наличии трендовой компоненты: наглядно выражена тенденция роста объемов продаж розничных магазинов. Средний абсолютный прирост для годовых уровней составил 126,2 млрд долл., т.е. в среднем ежегодно объем продаж в розничных магазинах США увеличивался на 126,2 млрд долл. В исследуемом периоде. На основе визуального анализа динамики данных можно предложить, что характер тенденции близок к линейному развитию.

На рис. 3 отчетливо видны сезонные колебания. Наблюдается устойчиво повторяющееся, резко «импульсное» увеличение объемов продаж в декабре каждого года (приближение новогодних праздников), характерны повторяющиеся из года в год сезонные «спады», приходящиеся на январь и февраль.

3)

Таблица 1 -**Объем продаж в розничных магазинах США (млрд долл.)**

Месяц	2011 г.	2012 г.	2013 г.	2014 г.
Январь	142,1	148,4	154,6	167,0
Февраль	143,1	145,0	155,8	164,0
Март	154,7	164,6	184,2	192,1
Апрель	159,1	170,3	181,8	187,5
Май	165,8	176,1	187,2	201,4
Июнь	164,6	175,7	190,1	202,6
Июль	166,0	177,7	185,8	194,9
Август	166,3	177,1	193,8	204,2
Сентябрь	160,6	171,1	185,9	192,8
Октябрь	168,7	176,4	189,7	194,0
Ноябрь	167,2	180,9	194,7	202,4
Декабрь	204,1	218,3	233,3	238,0

Так как амплитуда сезонных колебаний не изменяется с течением времени, носит примерно постоянный характер при наличии возрастающего тренда, то для описания и прогнозирования динамики временного ряда можно предложить модель с аддитивной сезонностью.

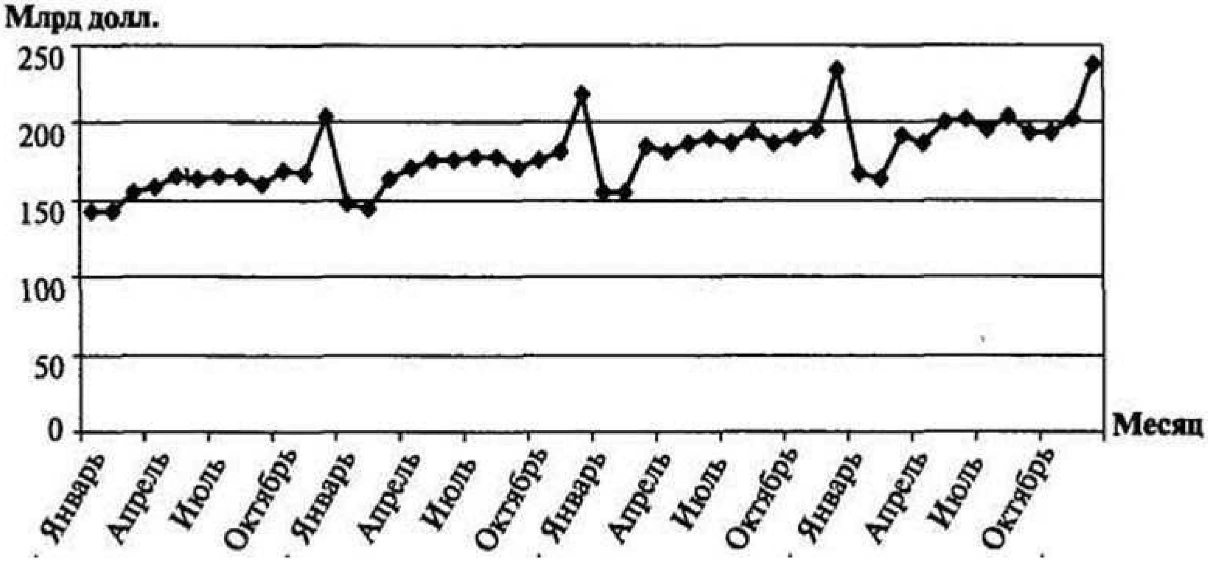


Рисунок 1 - Ежемесячная динамика объемов продаж в розничных магазинах

Результаты расчетов представлены в табл. 2.

Таблица 2 - **Прогнозирование объемов продаж в розничных магазинах**

t	Объем продаж в розничных магазинах y_t, y_t, млрд долл.	Скользящая средняя $y_t^{\prime}$	$x_t=y_t-y_t^{\prime}$		
1	2	3	4	5	6
1	142,1	-	-	165,517	132,405
2	143,1	-	-	169,178	133,629
3	154,7	-	-	156,385	158,907
4	159,1	-	-	162,017	158,560
5	165,8	-	-	161,191	166,971
6	164,6	-	-	159,717	168,130
7	166	163,788	2,213	164,384	165,748
8	166,3	164,129	2,171	162,753	168,564
9	160,6	164,621	-4,021	164,369	162,105
10	168,7	165,500	3200	167,677	167,810
11	167,2	166,369	0,804	164,399	170,473
12	204,1	167,288	36,813	164,688	207,969

13	148,4	168,238	-19,838	171,817	146,025
14	145	169,175	-24,175	171,078	144,248
15	164,6	170,063	-5,462	166,285	196,526
16	170,3	170,821	-0,521	173,217	169,179
17	176,1	171,713	4,388	171,491	177,591
18	175,7	172,875	2,825	170,817	178,749
19	177,7	173,725	3,975	176,084	176,368
20	177,1	174,433	2,667	173,553	179,183
21	171,1	175,700	-4,600	174,896	172,725
22	176,4	176,996	-0,596	175,377	178,430
23	180,9	177,938	2,963	178,099	181,092
24	218,3	179,000	39,300	178,88	218,589
25	154,6	179,938	-25,338	178,017	156,644
26	155,8	180,971	-25,171	181,878	154,868
27	184,2	182,283	1,917	185,885	180,146
28	181,8	183,454	-1,654	184,717	179,799
29	187,2	184,583	2,617	182,591	188,211
30	190,1	185,783	4,317	185,217	189,369
31	185,8	186,925	-1,125	184,184	186,988
32	193,8	187,783	6,017	190,253	189,803
33	185,9	188,454	-2,554	189,696	183,345
34	189,7	189,021	0,679	188,677	189,049
35	194,7	189,850	4,850	191,899	191,712
36	233,3	190,963	42,338	193,88	229,208
37	167	191,863	-24,863	190,417	167,264
38	164	192,675	-28,675	-28,675	165,488
39	192,1	193,396	-1,296	-1,296	190,766
40	187,5	193,863	-6,362	-6,362	190,419
41	201,4	194,363	7,038	7,038	198,830
42	202,6	194,879	7,721	7,721	199,989
43	194,9	-	-	-	197,607
44	204,2	-	-	-	200,423
45	192,8	-	-	-	193,965
46	194	-	-	-	199,669
47	202,4	-	-	-	202,332
48	238	-	-	-	239,828

Поясним последовательность вычислений. В гр. 3 табл. 2 показаны результаты сглаживания исходного временного ряда с помощью скользящей средней.

$$\hat{y} = \frac{\frac{1}{2}yt\text{-}6 + yt\text{-}5 + \cdots + yt + \cdots + yt{+}5 + \frac{1}{2}yt{+}6}{12}$$

$$y7' = \frac{\frac{1}{2} \cdot 142{,}1 + 143{,}1 + \cdots + 204{,}1 + \frac{1}{2} \cdot \cdot 148{,}4}{12} = 163{,}78_8$$

$$y8' = \frac{\frac{1}{2} \cdot 143{,}1 + 154{,}7 + \cdots + 148{,}4 + \frac{1}{2} 145}{12} = 164{,}12_9$$

В гр. 4 табл. 2 показаны уровни x_t, полученные вычитанием из исходного временного ряда значений сглаженного ряда. Уровни x_t отражают влияние случайных факторов и сезонности.

Предварительную оценку сезонности получим усреднением уровней временного ряда x_t для одноименных месяцев (табл 3)

$$\bar{x}_t = \frac{-19.838 - 25.338 - 24.86_3}{3} = -23.346$$

Таблица 3

Оценивание сезонной компоненты в аддитивной модели

Месяц	Порядковый номер i	Предварительная оценка сезонной компоненты $\bar{x}_i$	Скорректированные значения сезонной компоненты S_i
Январь	1	-23,346	-23,417
Февраль	2	-26,007	-26,078
Март	3	-1,614	-1,685
Апрель	4	-2,846	-2,917
Май	5	4,681	4,609
Июнь	6	4,954	4,883
Июль	7	1,687	1,616
Август	8	3,618	3,547
Сентябрь	9	-3,725	-3,796
Октябрь	10	1,094	1,023
Ноябрь	11	2,872	2,801
Декабрь	12	39,483	39,412
Сумма		0,853	0,000

Аналогичным образом вычислены предварительные оценки сезонной компоненты для остальных месяцев (табл. 4.6). Сумма этих оценок отлична от нуля:

$$\sum_{i=1}^{12} \bar{x}i = 0.853$$

Произведем корректировку значений сезонной составляющей. В соответствии с (4.4) найдем величину «поправки», на которую надо изменить предварительные оценки сезонности:

$$\bar{x} = \frac{0.853}{12} = 0.071$$

Скорректированные оценки сезонности приведены в последней графе табл. 6. Например, оценка сезонной компоненты для января рассчитываются следующим образом:

$$S_1=-23.346-0.071=-23.417.$$

Полученные оценки сезонности удовлетворяют свойству:

$$\sum_{i=1}^{12} Si = 0.$$

В табл. 2 (гр. 5-6) показаны основные этапы процесса прогнозирования по аддитивной модели. Сначала из исходного ряда удаляется сезонная составляющая, т.е. осуществляется десезонализация (сезонная корректировка) объемов продаж.

Для описания тенденции применяются модель линейного тренда, так как это согласуется с результатами графического анализа динамики показателя. Коэффициенты линейной модели определялись с помощью метода наименьших квадратов для временного ряда, представленного в гр. 5 табл. 4.

Построенная модель имеет вид:

$\hat{y}_t^{(1)}$=157.937+0.885t.

Расчетные уровни (гр. 6 табл. 2) были вычислены суммированием трендовых значений, полученных подстановкой последовательных значений времени t=1,2,…,48 в уравнение (4.6), были подставлены последовательные значения времени t 49,50,…,60.

Например, прогноз объема продаж в январе 2015 г. Определялся следующим образом:

$\hat{y}_{15}$ =157,937+0,885·49+(-23,417)=177,884 (млрд долл.).

Как следует из табл. 4.7 ожидаемый суммарный объем продаж в 1996г. Составил 2474 (млрд долл.).

На рис. 2 представлены фактические уровни исследуемого временного ряда и расчетные значения вместе с прогнозными оценками ежемесячных объемов продаж в 2015г.

О высокой точности модели на всем рассматриваемом времени интервале свидетельствует рассчитанное значение средней относительной ошибки по модулю, равное 1,33%.

Таблица 4 -**Результаты прогнозирования объемов продаж в розничных магазинах.. По аддитивной тренд – сезонной модели**

Месяц	t	Трендовая компонента	Сезонная компонента	Прогноз
Январь	49	201,301	-23,417	177,884
Февраль	50	202,186	-26,078	176,108
Март	51	203,071	-1,685	201,386
Апрель	52	203,956	-2,917	201,039
Май	53	204,841	4,609	209,450
Июнь	54	205,726	4,883	210,609
Июль	55	206,611	1,616	208,227
Август	56	207,496	3,547	211,043
Сентябрь	57	208,381	-3,796	204,585
Октябрь	58	209,266	1,023	210,289
Ноябрь	59	210,151	2,801	212,952
Декабрь	60	211,036	39,412	250,448
Сумма				2474,019

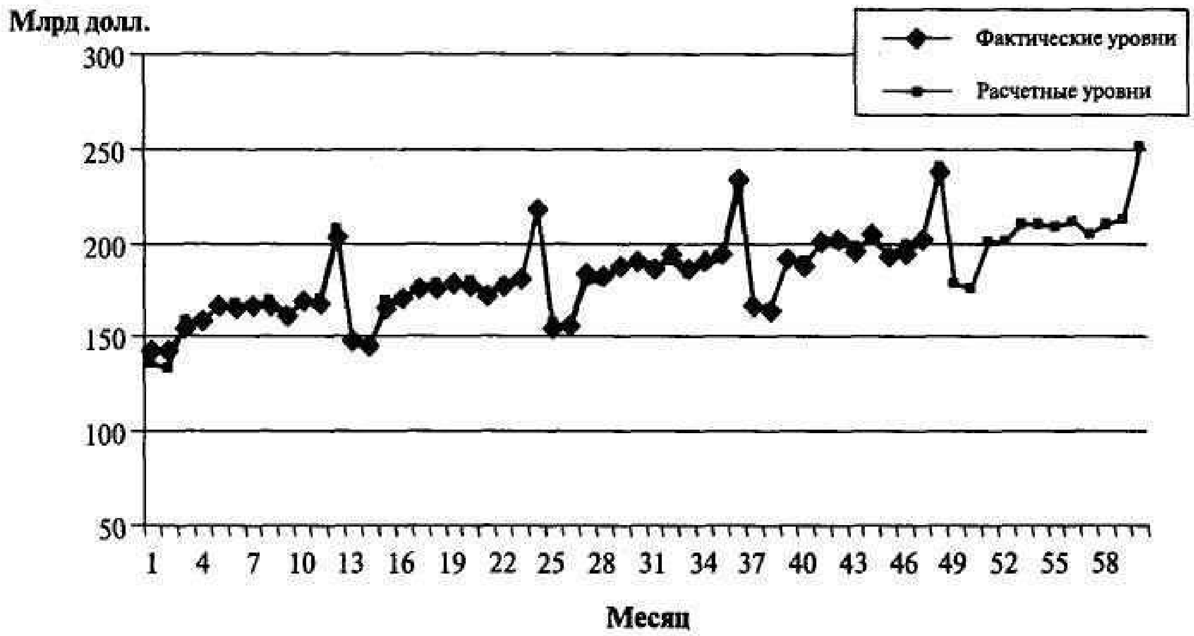


Рисунок 2 - Результаты прогнозирования объемов продаж в розничных магазинах

Также на практике достаточно часто для моделирования и прогнозирования сезонных колебаний могут быть использованы фиктивные переменны, получившие в англоязычной эконометрической литературе название dummy variables , манекены.

Контрольные вопросы

1. В динамике каких процессов проявляется влияние сезонного фактора?
2. Приведите примеры управленческих решений, принятие которых невозможно без учета сезонного фактора.
3. В каких случаях может применяться модель с аддитивной компонентой, в каких с мультипликативной?
4. Алгоритм десезонирования исходных данных?
5. Что показывает индекс сезонности?
6. Порядок прогнозирования с учетом сезонной компоненты временного ряда?

Лабораторная работа 5

Анализ и прогнозирование финансово-экономических показателей на основе моделей регрессии

Цель работы: Анализ и прогнозирование на основе моделей регрессии.

3.1. Основные теоретические сведения

4) ***Регрессионный анализ*** – позволяет определить аналитическое выражение связи между факторным признаком и результативным.

Математическую функцию, описывающую зависимость результативного признака Y от факторного X называют уравнением регрессии, а параметры модели – коэффициентами регрессии.

Для моделирования связи между признаками могут быть использованы любые математические функции. Чаще всего используются: полиномы n-порядка, степенная, показательная, гиперболическая.

Уравнение парной линейной регрессии имеет вид:

$\bar{y}_x = a_0 + a_1 x$, где

$\bar{y}_x$ - среднее значение результирующего признака У при при определенном значении факторного признака X;

a_0 - показывает усредненное влияние на результативный признак неучтенных в уравнении факторных признаков;

a_1 - коэффициент регрессии, показывает насколько в среднем изменится значение результативного признака при изменении факторного признака на единицу собственного измерения.

Оценка значение параметров уравнения регрессии осуществляется методом наименьших квадратов, при котором минимизируется сумма квадратов отклонений фактического значения результативного признака от теоретических, полученных по выбранному уравнению регрессии:

$$S = \sum (y - \overline{y}_x)^2 \to \min$$

Система нормальных уравнений для нахождения параметров линейной модели парной регрессии имеет вид:

$$\begin{cases} n \cdot a_0 + a_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \\ a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} x_i^{\,2} = \sum_{i=1}^{n} x_i \cdot y_i \end{cases}$$

, где n – объем выборки

Оценка качество модели регрессии

Качество модели регрессии связывают с адекватностью модели наблюдаемым (эмпирическим) данным. Проверка адекватности (или соответствия) модели регрессии наблюдаемым данным проводится на основе анализа остатков - ε_i.

После построения уравнения регрессии мы можем разбить значение *y* d каждом наблюдении на две составляющие :

Остаток представляет собой отклонение фактического значения зависимой переменной от значения данной переменной, полученное расчетным путем: .

Если , то для всех наблюдений фактические значения зависимой переменной совпадают с расчетными (теоретическими) значениями. Графически это означает, что теоретическая линия регрессии (линия, построенная по функции) проходит через все точки корреляционного поля, что возможно только при строго функциональной связи. Следовательно, результативный признак *y* полностью обусловлен влиянием фактора *x*.

На практике, как правило, имеет место некоторое рассеивание точек корреляционного поля относительно теоретической линии регрессии, т.е. отклонения эмпирических данных от теоретических ($\varepsilon_i \neq 0$). Величина этих отклонений и лежит в основе расчета показателей качества (адекватности) уравнения.

При анализе качества модели регрессии используется основное положение дисперсионного анализа [1], согласно которому общая сумма квадратов отклонений зависимой переменной от среднего значения может быть разложена на две составляющие - объясненную и необъясненную:

где - значения *y*, вычисленные по модели .

Коэффициент детерминации показывает какая доля вариации результативного признака, находящегося под воздействием изучаемых факторов, может быть объяснена построенным уравнением регрессии.

Коэффициент детерминации определяется следующим образом:

Чем ближе R^2 к 1, тем выше качество модели.

Также для оценки точности регрессионных моделей целесообразно использовать среднюю относительную ошибку аппроксимации:

Чем меньше рассеяние эмпирических точек вокруг теоретической линии регрессии, тем меньше средняя ошибка аппроксимации. Ошибка аппроксимации меньше 7% свидетельствует о хорошем качестве модели.

После того как уравнение регрессии построено, выполняется проверка значимости построенного уравнения в целом и отдельных параметров.

1. **Принятие решений на основе уравнения регрессии**

Полная экономическая интерпретация моделей регрессии позволяет выявить резервы развития субъектов рыночной экономики. Любая интерпретация начинается со статистической оценки уравнения регрессии в целом и оценки значимости входящих в модель факторных признаков.

Прежде всего, необходимо рассмотреть коэффициенты регрессии. Чем больше величина коэффициента регрессии, тем значительнее влияние данного признака на моделируемый.

Знаки коэффициентов регрессии говорят о характере влияния на результативный признак. Если факторный признак имеет знак плюс, то с увеличением данного фактора результативный признак возрастает; если факторный признак имеет знак минус, то с его увеличением результативный признак уменьшается.

Если экономические процессы подсказывают, что факторный признак должен иметь положительное значение, а он имеет знак минус, то необходимо проверить расчеты параметров уравнения регрессии. Такое явление чаще всего бывает в силу допущенных в расчетах ошибок.

Для того, чтобы на основании полученной модели сделать какие-то выводы для принятия решения или получить прогнозное значение некоторого параметра следует оценить значимость коэффициентов регрессии и адекватность всей модели.

Проверка значимости каждого коэффициента регрессии осуществляется с помощью t-критерия Стьюдента, расчетное значение которого определяется по формуле:

$$t_{\text{ðàñ÷}} = \frac{a_1}{\sqrt{\sigma^2_{a_i}}}$$, где

$\sigma^2_{a_i}$ - дисперсия коэффициента регрессии, наиболее простая формула которого имеет вид:

$$\sigma^2_{a_i} = \frac{\sigma^2_y}{k},$$ σ^2_y - дисперсия результирующего признака,

к- число факторных признаков в уравнении связи.

При проверки значимости коэффициентов регрессии на основе линейной парной зависимости дисперсия результирующего признака может быть вычислена следующим образом:

$$\sigma^2_{a_i} = \frac{\sigma^2_{\text{îñò}}}{\sigma^2_x (n-2)},$$ где

$$\sigma^2_x = \frac{\sum x_i^2}{n} - \overline{x}^2$$ - дисперсия факторного признака;

$$\sigma^2_{\text{îñò}} = \frac{\sum (y_i - \overline{y}_x)^2}{n}$$ - остаточная дисперсия.

Расчетное значение t-критерия Стьюдента сравнивается с критическим. которое определяется по таблице табулированных значений:

$$t_{\text{òàáë}} = \{\alpha; \nu = n - k - 1\}$$, где

α - уровень значимости критерия проверки гипотезы о равенстве нулю параметров уравнения регрессии;

ν - число степеней свободы, которое характеризует количество свободно варьируемых параметров совокупности;

к- количество параметров в уравнении регрессии.

Если расчетное значение t-критерия Стьюдента по модулю превышает табличное, то коэффициент регрессии признается значимым.

3.2. Задание на лабораторную работу

Для зависимой переменной *Y(t)* построить линейную модель регрессии, параметры которой оценить с помощью метода наименьших квадратов.

Оценить качество построенной модели (провести исследования адекватности и точности модели).

Рассчитать парный коэффициент корреляции переменных, коэффициент эластичности.

Отобразить на графике результаты аппроксимации и прогнозирования по модели регрессии, обосновать значимость прогноза.

3.3. Порядок выполнения работы

В таблице представлена информация об объемах продаж и затратах на рекламу одной фирмы, а также индекс потребительских расходов за ряд текущих лет. Требуется:

1. Построить диаграмму рассеяния (корреляционное поле) для переменных «объемы продаж» и «индекс потребительских расходов».

2. Определить степень влияния индекса потребительских расходов на объемы продаж (вычислить коэффициент парной корреляции).

3. Оценить значимость вычисленного коэффициента парной корреляции.

Таблица 1

Объем продаж, Y, тыс. руб.	126	137	148	191	274	370	432	445
Затраты на рекламу, X_1	4	4,8	3,8	8,7	8,2	9,7	14,7	18,7
Индекс потребительских расходов, X_2, %	100	98,4	101,2	103,5	104,1	107	107,4	108,5
Объем продаж, Y, тыс. руб.	367	367	321	307	331	345	364	384
Затраты на рекламу, X_1	19,8	10,6	8,6	6,5	12,6	6,5	5,8	5,7
Индекс потребительских расходов, X_2, %	108,3	109,2	110,1	110,7	110,3	111,8	112,3	112,9

Решение:

1. Вытянутость облака точек на диаграмме рассеяния вдоль наклонной прямой позволяет сделать предположение, что существует некоторая объективная тенденция прямой линейной связи между значениями переменных X — индекс потребительских расходов и Y — объемы продаж.

Рисунок 1- Диаграмма рассеяния (корреляционное поле)

2. Промежуточные расчеты при вычислении коэффициента корреляции между переменными X — индекс потребительских расходов и Y — объемы продаж приведены в табл. 2.

Средние значения случайных величин X и Y, которые являются наиболее простыми показателями, характеризующими последовательности $x_1, x_2, \ldots, x_{16}$ и $y_1, y_2, \ldots, y_{16}$, рассчитаем по формулам, соответственно

Дисперсия характеризует степень разброса значений
$x_1, x_2, \ldots, x_{16}$ ($y_1, y_2, \ldots, y_{16}$) вокруг своего среднего (и соответственно):

Стандартные ошибки случайных величин X и Y рассчитаем по формулам соответственно:

Коэффициент корреляции рассчитаем по формуле:

Таблица 2

№	Y	X					
1	2	3	4	5	6	7	8
1	126	100	-180,813	-7,231	1307,500	52,291	32693,160
2	137	98,4	-169,813	-8,831	1499,657	77,991	28836,285
3	148	101,2	-158,813	-6,031	957,838	36,376	25221,410
4	191	103,5	-115,813	-3,731	432,125	13,922	13412,535
5	274	104,1	-32,813	-3,131	102,744	9,805	1076,660
6	370	107	63,188	-0,231	-14,612	0,053	3992,660
7	432	107,4	125,188	0,169	21,125	0,028	15671,910
8	445	108,5	138,188	1,269	175,325	1,610	19095,785
9	367	108,3	60,188	1,069	64,325	1,142	3622,535
10	367	109,2	60,188	1,969	118,494	3,876	3622,535
1	2	3	4	5	6	7	8
12	307	110,7	0,188	3,469	0,650	12,032	0,035
13	331	110,3	24,188	3,069	74,225	9,417	585,035
14	345	111,8	38,188	4,569	174,469	20,873	1458,285
15	364	112,3	57,188	5,069	289,869	25,692	3270,410
16	384	112,9	77,188	5,669	437,557	32,135	5957,910
Сумма	4909	1715,7	0,000	0,000	**5681,994**	**305,474**	**158718,438**
Среднее	306,81	107,231					

3. Оценим значимость коэффициента корреляции. Для этого рассчитаем значение t-статистики по формуле

Табличное значение критерия Стьюдента равно: $t_{табл}$ ($\alpha = 0,1$; $k = n - 2 = 14$) = 1,76 (см. Приложение 2). Сравнивая числовые значения критериев, видно, что $t_{расч} > t_{табл}$, т.е. полученное значение коэффициента корреляции значимо. Таким образом, индекс потребительских расходов оказывает весьма высокое влияние на объемы продаж.

Построим линейную регрессионную модель методом наименьших квадратов. Используем промежуточные вычисления, приведенные в таблице 2:

Исходя из первых двух столбцов получим:

$$\mathrm{Y}_{ср} = \frac{4909}{16} = 306,81 \; ; \quad \mathrm{X}_{ср} = \frac{1715,7}{16} = 107,231 \; .$$

Остальные вычисления приведем в столбцах (3-8).

Вычислим значения коэффициентов линейной модели:

$$\mathrm{a}_1 = \frac{\sum(\mathbf{X_i} - \overline{\mathbf{X}})(\mathbf{Yi} - \overline{\mathbf{Y}})}{\sum(\mathbf{X_i} - \overline{\mathbf{X}})^2} = \frac{\mathbf{5681,994}}{\mathbf{305,474}} = 18,606 \; ,$$

$$\mathrm{a}_0 = \mathrm{Y}_{ср} - \mathrm{a}_1 \cdot \overline{\mathrm{X}} = \mathbf{306,81} + 18,606 \cdot \mathbf{107,231} = 2300,72$$

Таким образом, линейная модель имеет вид:

$$\overline{\mathrm{Y}}_{\mathrm{x}} = \mathrm{a}_0 + \mathrm{a}_1\mathrm{X} = 2300,72 + 18.606 \cdot \mathrm{X}$$

Таким образом, увеличение индекса потребительских расходов на 1% дает повышение объема реализации продукции на 18,6 тыс.рублей.

а) ***Коэффициент) детерминации)*** вычисляется по формуле:

$$D = R^2 = \frac{\sum_{t=1}^{T}(\hat{y}_t - \overline{y})^2}{\sum_{t=1}^{T}(y_t - \overline{y})^2} = 1 - \frac{\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{T}(y_t - \overline{y})^2}$$

$$\mathrm{D} = \mathrm{R}^2 = \frac{\mathbf{155544,069}}{\mathbf{158718,438}} = \mathbf{0,9885}$$

В случае линейной зависимости он показывает, какая часть общей дисперсии объясняется за счет вариации линейной комбинации независимых переменных $x_1, x_2, ..., x_n$ при данных коэффициентах $\alpha_0, \alpha_1, ..., \alpha_n$ регрессии.

D = 0,9885 ,т.е. 98,85% вариации объясняется факторами, включенными в уравнение регрессии, а 1,15% вариации объясняется прочими, не включенными в модель факторами.

b) **По критерию Фишера оценке качества всей модели**

$$F_{\text{ðàñ÷}} = \frac{\mathbf{r^2_{x,y}}}{\mathbf{1 - r^2_{x,y}}} \cdot (n-2) = \frac{0{,}816 \cdot 0{,}816}{1 - 0{,}816 \cdot 0{,}816} \cdot (16-2) = 26{,}76$$

F = 26,76

$F_{\text{табл.}} = 4{,}26$

$F > F_{\text{табл.}} \rightarrow$ гипотеза о заложенных в уравнении регрессии связей принимается в обеих моделях.

c) **Средняя относительная ошибка аппроксимации.**

Рассчитаем среднюю арифметическую величину относительной ошибки аппроксимации для каждой модели по формуле

$$E_{\text{отн}} = \frac{1}{T} \cdot \sum_{t=1}^{T} \left| \frac{(y_t - \hat{y}_t)}{y_t} \right| \cdot 100$$

Расчёты приведены в таблице. :

Получено $E_{\text{отн}} = 3{,}25$ %

Учитывая коэффициент детерминации, можно сделать вывод о том, что большая часть общей дисперсии объясняется за счет вариации линейной комбинации независимых переменной X при данных коэффициентах регрессии.

F-критерия Фишера (α =0,05) показывает существование заложенных в уравнении регрессии связей, что свидетельствует об адекватности и значимости уравнения регрессии. Средняя относительная ошибка аппроксимации (3,25%) достаточно мала, что говорит о достаточно высоком качестве построенной модели.

д) Дадим сравнительную оценку силы связи фактора с результатом с помощью коэффициента эластичности.

Для нахождения средних по совокупности показателей эластичности получаем формулу:

$$\overline{Э_{yx_i}} = \alpha_i \frac{\overline{x_i}}{\overline{\overline{y_{x_i}}}} \qquad \overline{Ý_{yx_1}} = 18{,}606 \cdot \frac{\mathbf{306{,}81}}{\mathbf{107{,}231}} \cdot \mathbf{100} = 6{,}49$$

Коэффициент эластичности показывает, что с ростом потребительских расходов на 1% реализации продукции возрастёт на 6,49%.

Коэффициент эластичности указывают на наличие незначительной связи фактора с результатом.

3.4. Варианты заданий

Номер Вашего варианта выбирается в соответствии с номером в списке группы. В соответствии с ним из таблицы выберите показатель *Y(t),* а данные фактора *X(t)* возьмите из следующей по порядку строки.

3.5. Контрольные вопросы

1. Какие ограничения накладываются на количество факторов, включаемых в регрессионную модель и в чем они вызваны?

2.Какими средствами оценивается качество построенных регрессионных моделей?

3. Что показывает коэффициент детерминации.

4. Каким образом на основе регрессионной модели получается прогноз зависимой переменной?

3.6 Список рекомендуемой литературы

1. Прогнозирование и планирование в условиях рынка [Текст] : учебное пособие / Т. Н. Бабич [и др.] - Москва : ИНФРА-М, 2013. - 336 с.

2. Дуброва, Т. А. Прогнозирование социально-экономических процессов [Текст] : учебное пособие / Т. А. Дуброва. - 2-е изд., испр. и доп. - М. : Маркет ДС, 2010. - 192 с.

3. Вертакова, Ю. В. Прогнозирование и планирование в условиях рынка [Текст] : учебное пособие / Ю. В. Вертакова, И. А. Козьева ; Курский государственный технический университет. – Курск : КурскГТУ, 2009. - 192 с.

4. Цыгичко, В. Н. Прогнозирование социально-экономических процессов [Текст] / В. Н. Цыгичко ; предисл. Д. М. Гвишиани. - 3-е изд. - М. :Либроком, 2009. - 238 с.

5. Кузык, Б. Н. Прогнозирование, стратегическое планирование и национальное программирование [Текст] : учебник / Б. Н. Кузык, Ю. В. Яковец, В. И. Кушлин. - 3-е изд., перераб. и доп. - М. : Экономика, 2009. - 591 с.