Т.И. Белых, А.В. Бурдуковская

ПРЕДВАРИТЕЛЬНЫЙ АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

Учебное пособие

Министерство образования и науки Российской Федерации
Байкальский государственный университет

Т.И. Белых, А.В. Бурдуковская

ПРЕДВАРИТЕЛЬНЫЙ АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

Учебное пособие

Иркутск
Издательство БГУ
2018

УДК 512.64(075.8)
ББК 22.143я7
Л59

Печатается по решению редакционно-издательского совета
Байкальского государственного университета

Рецензенты канд. физ.-мат. наук, доц. В.В. Братищенко
канд. экон. наук, доц. Т.И. Хитрова

Белых Т.И.
Л59 Предварительный анализ временных рядов [Электронный ресурс] : учеб. пособие / Т.И. Белых, А.В. Бурдуковская. – Иркутск: Изд-во БГУ, 2018. – 105 с. – Режим доступа: lib-catalog@bgu.ru.

В учебном пособии излагаются основные вопросы, включенные в программу дисциплины «Модели и методы прогнозирования», приводятся необходимые теоретические сведения, а также подробно разбираются типовые задачи и примеры, приводятся варианты индивидуальных заданий.

Учебное пособие предназначено для студентов 3, 4 курса всех форм обучения специальностей 09.03.03 Прикладная информатика, 38.03.05 Бизнес-информатика, может быть использовано как развернутый справочник для успешного усвоения дисциплины, систематизации и углублению знаний, и привитию навыков решения различных классов задач.

УДК 512.64(075.8)
ББК 22.143я7

Оглавление

Введение

Динамические процессы, происходящие в экономических системах, чаще всего проявляются в виде ряда последовательно расположенных в хронологическом порядке значений того или иного показателя, который в своих изменениях отражает ход развития изучаемого явления в экономике.

Последовательность наблюдений одного показателя, упорядоченных в зависимости от последовательно возрастающих или убывающих значений другого показателя, называют динамическим рядом, или рядом динамики. Если в качестве признака, в зависимости от которого происходит упорядочение, берется время, то такой динамический ряд называется временным рядом.

Так как в экономических процессах, как правило, упорядочение происходит в соответствии со временем, то при изучении последовательных наблюдений экономических показателей все три приведенных выше термина используются как равнозначные. Составными элементами рядов динамики являются, таким образом, цифровые значения показателя, называемые уровнями этих рядов, и моменты или интервалы времени, к которым относятся уровни.

Временные ряды, образованные показателями, характеризующими экономическое явление на определенные моменты времени, называются моментными; если уровни временного ряда образуются путем агрегирования за определенный промежуток (интервал) времени, то такие ряды называются интервальными временными рядами.

Временные ряды могут быть образованы как из абсолютных значений экономических показателей, так и из средних или относительных величин.

Под длиной временного ряда понимают время, прошедшее от начального момента наблюдения до конечного. Часто длиной ряда называют количество уровней, входящих во временной ряд.

Если во временном ряду проявляется длительная («вековая») тенденция изменения экономического показателя, то говорят, что имеет место тренд. Таким образом, под трендом понимается изменение, определяющее общее направление развития, основную тенденцию временных рядов. В связи с этим экономико-математическая динамическая модель, в которой развитие моделируемой экономической системы отражается через тренд ее основных показателей, называется трендовой моделью.

Отличие временных экономических рядов от простых статистических совокупностей заключается, прежде всего, в том, что последовательные значения уровней временного ряда зависят друг от друга. Поэтому применение выводов и формул теории вероятностей и математической статистики требует известной осторожности при анализе временных рядов, особенно при экономической интерпретации результатов анализа.

Исследование временных рядов осуществляется с различными целями, поэтому применяемые подходы и соответствующие математические модели зависят от поставленных задач.

Существуют две основные цели анализа временных рядов: определение природы ряда и прогнозирование (предсказание будущих значений временного

ряда по настоящим и прошлым значениям). Обе эти цели требуют, чтобы модель ряда была идентифицирована и, более или менее, формально описана. Как только модель определена, вы можете с ее помощью интерпретировать рассматриваемые данные (например, использовать в вашей теории для понимания сезонного изменения цен на товары, если занимаетесь экономикой). Не обращая внимания на глубину понимания и справедливость теории, вы можете экстраполировать затем ряд на основе найденной модели, т.е. предсказать его будущие значения.

Таким образом, анализ временных рядов состоит в построении по возможности простых моделей, адекватно описывающих имеющиеся ряды наблюдений и составляющих базу для решения, в первую очередь, следующих задач: объяснение механизма формирования уровней ряда, построение прогноза будущих значений временного ряда.

Предварительная обработка временных рядов состоит в выявлении аномальных значений ряда и сглаживании ряда. Аномальные значения временного ряда не отвечают потенциалу исследуемой экономической системы, и их использование для построения трендовой модели может сильно исказить получаемые результаты. Причинами появления аномальных уровней могут быть технические ошибки при сборе, обработке и передаче информации. Такие ошибки называются ошибками первого рода, их можно выявить и устранить или принять меры к их недопущению. Кроме того, аномальные уровни могут возникать из-за воздействия факторов, имеющих объективный характер, но действующих эпизодически. Такие ошибки называются ошибками второго рода, их невозможно устранить, но можно исключить из рассмотрения, заменив аномальное значение на среднеарифметическое двух соседних уровней.

Очень часто уровни экономического ряда динамики колеблются, так что тенденция развития экономического процесса скрыта случайными отклонениями. Сглаживание временного ряда позволяет отфильтровать мелкие случайные колебания и выявить основную тенденцию изменения исследуемой величины. При механическом сглаживании выравнивание отдельных уровней производится с использованием значений соседних уровней

Выявление аномальных уровней временного ряда

Пусть дан временной ряд $y_t, t = \overline{1,n}$. Одним из методов выявления аномальных уровней ряда является метод Ирвина, согласно которому аномальной считается точка y_t, отстоящая от предыдущей точки y_{t-1} на величину, большую среднеквадратичного отклонения

Технология применения метода Ирвина:

1) рассчитать значения параметра λ_t по формуле:

$$\lambda_t = \frac{|y_t - y_{t-1}|}{\sigma_y}, \quad t = \overline{2,n} \quad ,$$

где

n – общее число уровней ряда;

y_t, y_{t-1} – соответственно текущее и предыдущее значения уровней ряда;

σ_y – среднее квадратическое (стандартное) отклонение показателя

$$\sigma_y = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \bar{y})^2}{n-1}},$$

$\bar{y}$ – среднее значение показателя

$$\bar{y} = \frac{\sum_{t=1}^{n} y_t}{n} \quad ;$$

2) сравнить расчетные значения λ_t с табличным значением критерия Ирвина λ_α (таблица 1):

Таблица 1

Критические значения распределения Ирвина

n	2	3	10	20	30	50	100
λ_α	2,8	2,3	1,5	1,3	1,2	1,1	1

Табличное значение λ_α определяется при уровне значимости α и числе степеней свободы k. Уровень значимости α, для экономических расчетов, принимается равным 0,05 или 0,01; число степеней свободы равно n-2.

Если $\lambda_t > \lambda_\alpha$, то соответствующее значение y_t уровня ряда считается аномальным. Результаты представить в виде таблицы 2:

Таблица 2

Оформление работы

t	y_t	λ_t	*Сравнение λ_t с λ_α*
1			
2			
…			
n			

Примечание 1: столбец «*Сравнение* λ_t *с* λ_α» содержит значения «аномальный», «не аномальный».

Пример 1. Задан временной ряд y_t, t=1,2, …, 15 (таб. 3). Выявить аномальные уровни ряда.

1 шаг. Рассчитать стандартное отклонение временного ряда σ_y с помощью функции СТАНДОТКЛОН().

2 шаг. Рассчитать λ_t для всех уровней, начиная со второго.

3 шаг. Определить по таблице 1 критическое значение и сделать выводы.

Таблица 3

Исходные данные для выявления аномальных уровней временного ряда

y_t	λ_t	*Сравнение* λ_t *с* λ_α
87		
96	0,181521	не аномальный
107	0,22186	не аномальный
119	0,242029	не аномальный
129	0,201691	не аномальный
143	0,282367	не аномальный
156	0,262198	не аномальный
167	0,22186	не аномальный
176	0,181521	не аномальный
186	0,201691	не аномальный
198	0,242029	не аномальный
211	0,262198	не аномальный
220	0,181521	не аномальный
228	0,161352	не аномальный
239	0,22186	не аномальный
σ_y =	49,58091	
λ_α =	1,5	

Вывод: аномальные уровни временного ряда – отсутствуют (таб. 3).

Замечание 1. В случае присутствия аномальных уровней ряда следует определить причины их возникновения. Если точно установлено, что они вызваны ошибками первого рода, то они устраняются либо заменой простой средней арифметической двух соседних уровней ряда, либо заменой соответствующими значениями по кривой, аппроксимирующей данный ряд.

Методы выявления наличия тренда

Первая задача, которая возникает при анализе рядов динамики, заключается в выявлении и описании основной тенденции развития изучаемого явления (тренда). Трендом называется плавное и устойчивое изменение уровней явления во времени, свободное от случайных колебаний.

Изучение тренда включает в себя два этапа:

1. проверка ряда на наличие тренда;

2. выравнивание ряда динамики и непосредственное выделение тренда.

Проверка ряда на наличие тренда проводится разными методами, самым простым из которых является метод средних.

Метод разности средних уровней

Суть метода заключается в следующем: изучаемый ряд динамики разбивается на несколько интервалов (чаще всего на два), для каждого из которых определяется средняя величина - $\overline{y}_1$ и $\overline{y}_2$. Выдвигается гипотеза о существенном различии средних. Если выдвинутая гипотеза принимается, то признается наличие тренда.

Рекомендуется следующая последовательность действий:

1 шаг. Исходный ряд $y_t, t = \overline{1,n}$ разбить на две приблизительно равные части, т.е. $n_1 \approx n_2$, $n_1+n_2 = n$;

2 шаг. Для каждой части вычислить средние значение уровней показателя $\overline{y_1}$ *и* $\overline{y_2}$:

$$\overline{y}_1 = \frac{\sum_{t=1}^{n_1} y_t}{n_1}, \quad \overline{y}_2 = \frac{\sum_{t=n_1+1}^{n_2} y_t}{n_2}$$

и дисперсии D_1 *и* D_2:

$$D_1 = \frac{\sum_{t=1}^{n_1} (y_t - \overline{y}_1)^2}{n_1 - 1}, \quad D_2 = \frac{\sum_{t=n_1+1}^{n_2} (y_t - \overline{y}_2)^2}{n_2 - 1};$$

3 шаг. Проверить гипотезу о равенстве дисперсий обеих частей ряда

$$H_0: D_1=D_2$$

$$H_1: D_1 \neq D_2$$

с помощью критерия Фишера. Для этого вычислить расчетное значение критерия Фишера ($F_{расч}$) и сравнить его с табличным значением F-критерия ($F_{табл}$ выбирается при уровне значимости $\alpha = 0{,}05$, реже 0,01, и степенях свободы $n_1 - 1$; $n_2 - 1$).

Правило вычисление расчетного значения критерия Фишера:

$$F_{расч} = \begin{cases} D_1 / D_2, & если \quad D_1 > D_2 \\ D_2 / D_1, & если \quad D_1 < D_2 \end{cases}.$$

Если $F_{расч} > F_{табл}$, то гипотеза отвергается и делается вывод, что данный метод для определения наличия тренда ответа не дает и необходимо применение какого-либо другого метода. Если $F_{расч} < F_{табл}$, то гипотеза о равенстве дисперсий обеих частей ряда принимается;

2) проверить гипотезу об отсутствии тренда

$$H_0 : \overline{y_1} = \overline{y_2}$$
$$H_1 : \overline{y_1} \neq \overline{y_2}$$

с использованием t- критерия Стьюдента

$$t_{расч} = \frac{\overline{y_1} - \overline{y_2}}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\sigma} \quad ,$$

где σ - среднее квадратическое отклонение разности средних:

$$\sigma = \sqrt{\frac{(n_1 - 1)D_1 + (n_2 - 1)D_2}{n_1 + n_2 - 2}} .$$

Если $t_{расч} > t_{табл}$ (табличное значение принимается при уровне значимости 0,05, реже 0,01), то делается вывод о наличии тренда.

Замечание 2. Рассмотренный выше метод в ряде случаев дает вполне приемлемые результаты. Однако следует отметить, что ему свойственны весьма существенные дефекты. Прежде всего, он применим только для рядов с монотонной тенденцией. Если же ряд меняет общее направление развития, то точка поворота тенденции может оказаться близкой к середине ряда, в силу этого средние двух отрезков ряда будут близки и проверка может не показать наличие тренда. Вместе с тем можно выдвинуть и более серьезное возражение, основанное на том, что величина среднего квадратического отклонения, с которой сравнивается разность средних, зависит в динамическом ряду не только от колеблемости уровней, но и от самого тренда. Иначе говоря, существование тренда сказывается на показателе среднего квадратического отклонения. Сама же разность средних в значительной мере будет определяться тем, какой угол наклона имеет тренд.

Пример 2. Имеется временной ряд y_t, $t = \overline{1, 49}$, содержащий данные об индексе потребительских цен (в % к предыдущему месяцу) платных услуг (таб. 4). Требуется определить наличие тренда методом средних разностей.

Таблица 4

Исходные данные для выявления наличия тренда методом средних разностей

№	y_t	№	y_t	№	y_t	№	y_t	№	y_t
1	123	12	103	23	86	34	69	45	51
2	119	13	100	24	84	35	70	46	46
3	110	14	102	25	85	36	65	47	42
4	112	15	103	26	83	37	63	48	43
5	111	16	99	27	84	38	67	49	41
6	109	17	98	28	82	39	62		
7	110	18	102	29	79	40	57		
8	108	19	94	30	74	41	59		
9	107	20	87	31	76	42	53		
10	109	21	89	32	73	43	48		
11	107	22	100	33	74	44	50		

1 шаг. Разбить временной ряд на 2 приблизительно равные подвыборки n_1=24, n_2=25.

2 шаг. Найти средние значения и дисперсии с помощью описательной статистики ДАННЫЕ/АНАЛИЗ ДАННЫХ/ОПИСАТЕЛЬНАЯ СТАТИСТИКА (рис. 1):

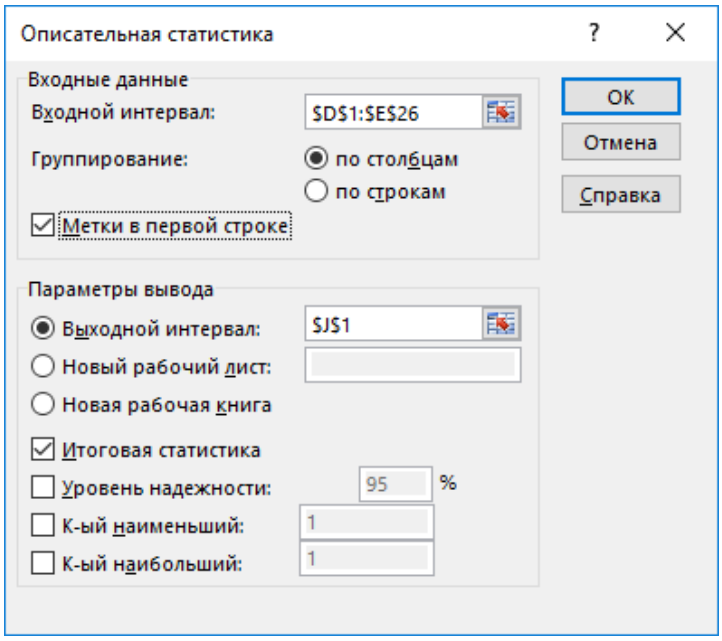


Рис. 1. Диалоговое окно Описательная статистика

Описательная статистика выводит таблицу, в которой рассчитываются статистические характеристики выборок (таб. 5).

Таблица 5

Результат описательной статистики

Выборка 1		*Выборка 2*	
Среднее	103	**Среднее**	63,84
Стандартная ошибка	2,028778	Стандартная ошибка	2,84176
Медиана	103	Медиана	65
Мода	110	Мода	74
Стандартное отклонение	9,938944	Стандартное отклонение	14,2088
Дисперсия выборки	98,78261	**Дисперсия выборки**	201,89
Эксцесс	-0,13601	Эксцесс	-1,25168
Асимметричность	-0,23566	Асимметричность	-0,13063
Интервал	39	Интервал	44
Минимум	84	Минимум	41
Максимум	123	Максимум	85
Сумма	2472	Сумма	1596
Счет	24	Счет	25

3 шаг. Рассчитать среднее квадратическое отклонение разности средних σ, t–статистику и F–статистику. Сравнить с критическими значениями.

□=	12,30582			
$t_{расч}$	11,1355	>	$t_{крит}$	2,011741
$F_{расч}$	2,043781	>	$F_{крит}$	1,993239

При проверке гипотезы о равенстве средних предполагается, что дисперсии двух совокупностей незначительно различаются между собой. Проверка однородности дисперсий реализуется с помощью F − статистики Фишера, который основан на сравнении расчетного отношения с табличным. Если расчетное значение *F* меньше, чем табличное, при заданном уровне вероятности, то можно принять гипотезу о равенстве дисперсий. Если же *F* больше, чем табличное значение, то гипотеза о равенстве дисперсий отклоняется.

Вывод. Можно сделать вывод о том, что данный метод для определения наличия тренда ответа не дает и необходимо применение какого-либо другого метода

Выявление наличия тренда методом разности средних уровней средствами Анализа данных

1 способ.

1 шаг. На рабочем листе сформировать исходный ряд (рекомендуется разместить значения уровней ряда в столбец);

2 шаг. Выполнить команду **ДАННЫЕ/АНАЛИЗ ДАННЫХ/Двухвыборочный F-тест для дисперсии**;

3 шаг. В открывшемся диалоговом окне (рис. 2) «Двухвыборочный F-тест для дисперсии» задать параметры:

– в поле *«Интервал переменной 1»* – ссылка на диапазон, содержащий первую часть ряда;

– в поле *«Интервал переменной 2»* – ссылка на диапазон, содержащий вторую часть ряда;

– в поле *«Альфа»* – требуемый уровень значимости;

– для вывода результатов задать параметры вывода: или новый рабочий лист, когда результаты выводятся с ячейки A1 или незаполненная ячейка текущего рабочего листа, начиная с которой выводится результат.

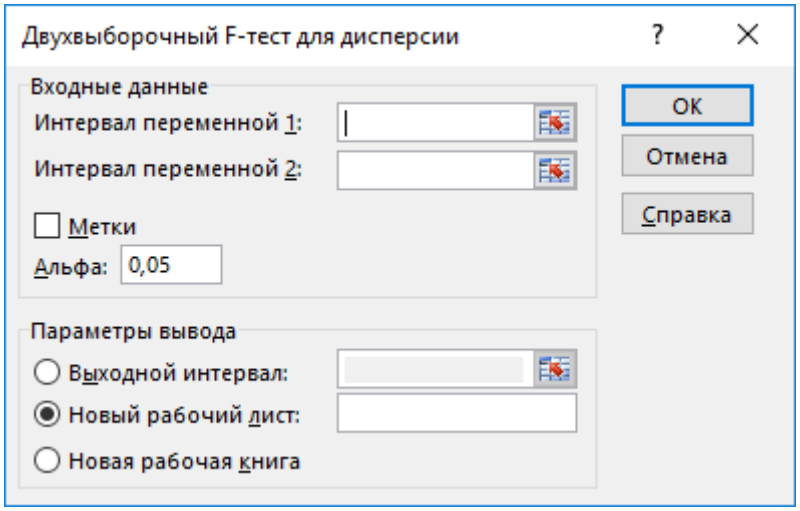


Рис. 2. Диалоговое окно двухвыборочного F-теста для дисперсии

– нажать кнопку ОК.

Общее описание результата проведения двухвыборочного F-теста для дисперсии (таб. 6).

Если $F_{расч} < F_{табл}$, то гипотеза о равенстве дисперсий обеих частей ряда принимается и делается вывод о наличии тренда (в противном случае делается вывод о невозможности определения наличия тренда данным методом и следует воспользоваться другим).

Таблица 6

Общий вид результатов проведения двухвыборочного F-теста для дисперсии

	Переменная 1	*Переменная 2*
Среднее	$\bar{y}_1$	$\bar{y}_2$
Дисперсия	D_1	D_2
Наблюдения	n_1	n_2
df (число степеней свободы)	$n_1 - 1$	$n_2 - 1$
F	$F_{расч}$	
P(F<=f) одностороннее		
F критическое одностороннее	$F_{табл}$	

Таблица 7

Результаты проведения двухвыборочного F-теста для дисперсии

Двухвыборочный F-тест для дисперсии		
	Выборка 1	*Выборка 2*
Среднее	103	63,84
Дисперсия	98,7826087	201,89
Наблюдения	24	25
df	23	24
F	0,48928926	
P(F<=f) одностороннее	0,045589095	
F критическое одностороннее	0,498750764	

2 способ.

1 шаг. Выполнить команду **СЕРВИС/АНАЛИЗ ДАННЫХ/Двухвыборочный t-тест с одинаковыми дисперсиями;**

2 шаг. В открывшемся диалоговом окне (рис. 3) «Двухвыборочный t-тест с одинаковыми дисперсиями» задать параметры:

– в поле *«Интервал переменной 1»* – ссылка на диапазон, содержащий первую часть ряда;

– в поле *«Интервал переменной 2»* – ссылка на диапазон, содержащий вторую часть ряда;

– в поле *«Гипотетическая средняя разность»* – значение 0;

– в поле «Альфа» – требуемый уровень значимости;

– для вывода результатов рекомендуется использовать поле «Новый рабочий лист».

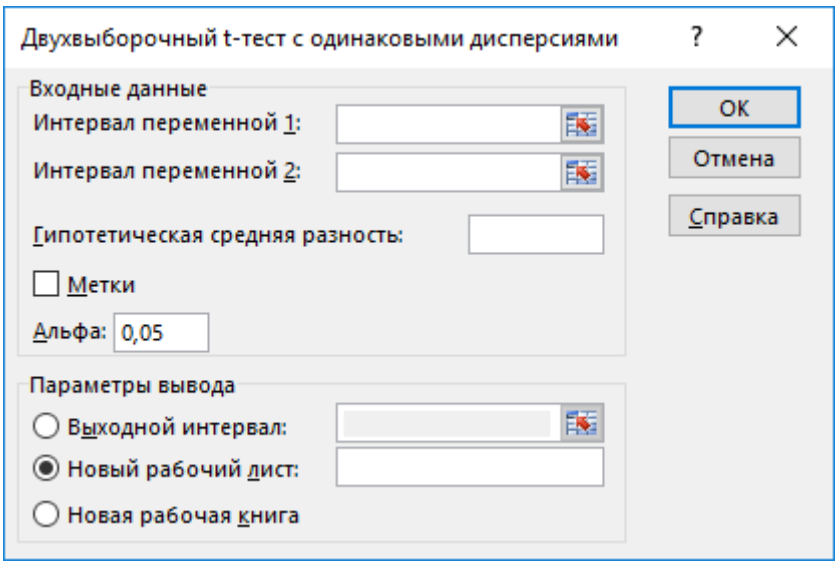


Рис. 3. Двухвыборочный t-тест с одинаковыми дисперсиями

3 шаг. Для получения результатов нажать кнопку ОК.

Результатом выполнения данной команды является таблица 8. Приведем общее описание результата проведения двухвыборочного F-теста с одинаковыми дисперсиями (таблица 9).

Так как $t_{расч} > t_{табл}$, то делаем вывод о наличии тренда во временном ряду.

Таблица 8

Общий вид результатов проведения двухвыборочного t-теста с одинаковыми дисперсиями

	Переменная 1	Переменная 2
Среднее	$\bar{y}_1$	$\bar{y}_2$
Дисперсия	D_1	D_2
Наблюдения	n_1	n_2
Объединенная дисперсия	σ	
Гипотетическая разность средних		
Df	n - 2	
t-статистика	$t_{расч}$	
P(T<=t) одностороннее		
t критическое одностороннее		
P(T<=t) двухстороннее		
t критическое двухстороннее	$t_{табл}$	

Таблица 9

Результаты проведения двухвыборочного t-теста с одинаковыми дисперсиями

Двухвыборочный t-тест с одинаковыми дисперсиями		
	Выборка 1	*Выборка 2*
Среднее	103	63,84
Дисперсия	98,78261	201,89
Наблюдения	24	25
Объединенная дисперсия	151,4332	
Гипотетическая разность средних	0	
df	47	
t-статистика	11,1355	
P(T<=t) одностороннее	4,45E-15	
t критическое одностороннее	1,677927	
P(T<=t) двухстороннее	8,89E-15	
t критическое двухстороннее	2,011741	

Метод Фостера-Стюарта

Проверка разности средних не является единственным способом проверки гипотез о наличии тренда средней динамического ряда. Метод разработанный Ф. Фостером и А. Стюартом дает более надежные результаты, чем остальные. Этот метод позволяет установить наличие тренда не только у временного ряда, но и выявить его в значении дисперсии уровней, что немаловажно для прогностического анализа. При отсутствии тренда дисперсии разброс уровней ряда постоянен, при наличии тренда дисперсии дисперсия увеличивается или уменьшается.

Ф. Фостер и А. Стюарт предложили по данным исследуемого ряда определять величины k_t и l_t. Значения k_t и l_t находятся путем последовательного сравнения уровней. Если какой-либо уровень ряда превышает по своей величине каждый из предыдущих уровней, то величине k_t присваивается значение 1, в остальных случаях она равна 0. Таким образом,

Метод Фостера-Стюарта рекомендуется выполнять по следующим этапам:

1 шаг. Сравнить каждый уровень исходного ряда $y_t, t = \overline{1,n}$, начиная со второго, со всеми предыдущими уровнями, и построить две числовые последовательности:

$$k_t = \begin{cases} 1, & \text{если } y_t > \text{всех предыдущих уровней} \\ 0, & \text{в противном случае} \end{cases},$$

$$l_t = \begin{cases} 1, & \text{если } y_t < \text{всех предыдущих уровней} \\ 0, & \text{в противном случае} \end{cases} \qquad t = \overline{2,n};$$

2 шаг. Вычислить величины S и d:

$$S = \sum_{t=2}^{n}(k_t + l_t) \ , \qquad d = \sum_{t=2}^{n}(k_t - l_t) \ ,$$

Нетрудно видеть, что k_t принимает значения 0 и 1: $k_t = 0$ в случае, если y_t не является ни наибольшим, ни наименьшим уровнем среди всех предшествующих уровней, в противном случае $k_t = 1$. Легко определить, что S может находиться в пределах $0 \le S \le n - 1$. (Здесь, как и выше, n означает число членов ряда.) Если все уровни равны (нулевая дисперсия), то $S = 0$, если же они монотонно растут, или падают, или колебания их чередуются, систематически увеличиваясь или падая, то $S = n - 1$.

В свою очередь величина l_t принимает значения 0; 1 и –1. Найдем теперь пределы для d: нижний предел равен $-(n - 1)$, верхний составляет $n - 1$. Нижний предел соответствует монотонно убывающему, а верхний – монотонно растущему ряду.

Показатели S и d асимптотически нормальны и имеют независимые распределения. Они существенно зависят от порядка расположения уровней во времени. Показатель S применяется для обнаружения тенденций изменения дисперсии, d – для обнаружения тенденций в средней. После того как для исследуемого ряда найдены фактические значения d и S, проверяется гипотеза о том, можно ли

считать случайными разности $d - 0$ и $S - \mu$. Величина S характеризует изменение временного ряда, она может принимать значение от 0 (когда все уровни ряда равны) до $n - 1$ (ряд монотонный). Величина d характеризует изменение дисперсии временного ряда и изменяется от $-(n - 1)$ (когда ряд монотонно убывает) до $(n - 1)$ (когда ряд монотонно возрастает). Эти величины являются случайными с математическим ожиданием μ для значения s и 0 для значения d.

Результаты вычислений рекомендуется представить в следующем виде (таб. 10):

Таблица 10

Оформление результатов определения наличия тренда методом Фостера-Стюарта

t	y_t	k_t	l_t	$k_t + l_t$	$k_t - l_t$
1					
2					
…					
n					
Итого				***S***	***d***

3 шаг. Проверить гипотезы.

Первая гипотеза: можно ли считать случайным отклонение S от его математического ожидания μ:

$$H_0: S = \mu$$
$$H_1: S \neq \mu.$$

Для этого:

– рассчитать значение t-критерия по формуле $t_S = \frac{|S - \mu|}{\sigma_s}$,

где σ_s – стандартное отклонение S рассчитывается по формуле

$$\sigma_s = \sqrt{2\ln(n) - 3{,}4253} \quad ,$$

μ – табличное значение математического ожидания:

n	10	20	30	40
μ	3,858	5,195	5,99	6,557

– расчетное значения t_s сравнить с табличным значением $t_{табл}$ (при уровне значимости 0,05 и степени свободы $k = n-2$). В случае, если $t_s > t_{табл}$, то для данного ряда имеется тренд, т.е. гипотеза об отсутствии тренда отвергается;

Вторая гипотеза: можно ли считать случайным отклонение d от нуля?

$$H_0: d = 0$$
$$H_1: d \neq 0.$$

Для этого:

– рассчитать значение t-критерия по формуле $t_d = |d| / \sigma_d$, где σ_d - стандартное отклонение d вычисляемое по формуле

$$\sigma_d = \sqrt{2\ln(n) - 0{,}8456} \quad ;$$

– расчетное значения t_d сравнить с табличным значением $t_{табл}$ (его целесообразно при уровне значимости 0,05 и степени свободы *n-2*). Если $t_d < t_{табл}$, то для данного ряда тренда дисперсии уровней ряда нет, и гипотеза H_0 принимается.

Результаты проведения исследования целесообразно представить в виде (таб. 11):

Таблица 11

Пример оформления расчетов

S	
d	
Стандартное отклонение S	
Стандартное отклонение d	
Критерий Стьюдента для S	
Критерий Стьюдента для d	
Математическое ожидание	
Табличный критерий Стьюдента	

Пример 3. Имеется временной ряд y_t, содержащий данные об объеме перевозок грузов железнодорожным транспортом, млн. тонн.

y_t	200	310	320	260	190	210	310	410	430	370	300	320	340	350	410	435

Требуется определить наличие тренда методом Фостера-Стюарта.

1 шаг. Составить последовательности k_t и l_t, рассчитать сумму $k_t + l_t$ и разность $k_t - l_t$. Подсчитать итоговые суммы *S* и *d (таб. 12)*.

Таблица 12

Исходные данные примера 3

y_t	k_t	l_t	$kt+l_t$	$kt-l_t$
200				
310	1	0	1	1
320	1	0	1	1
260	0	0	0	0
190	0	1	1	-1
210	0	0	0	0
310	0	0	0	0
410	1	0	1	1
430	1	0	1	1
370	1	0	1	1
300	0	0	0	0

y_t	k_t	l_t	$kt+l_t$	$kt-l_t$
320	0	0	0	0
340	0	0	0	0
350	0	0	0	0
410	0	0	0	0
435	1	0	1	1
Итого			7	5

2 шаг. Рассчитать стандартные отклонения и t – статистики для последовательностей (рис. 4).

S	7
d	5
Стандартное отклонение S	1,45597989
Стандартное отклонение d	2,16785088
Критерий Стьюдента для S	2,15799684
Критерий Стьюдента для d	2,3064317
Математическое ожидание	3,858
Табличный критерий Стьюдента	2,14478669

Рис. 4. Пример расчета стандартных отклонений и t-статистик

3 шаг. Сравнить значения: значения t-статистик, t_s и t_d больше критического значения, что подтверждает гипотезу о наличии тренда.

Критерий «восходящих» и «нисходящих» серий

Этот критерий улавливает постепенное смещение среднего значения в исследуемом распределении не только монотонного, но и более общего, например, периодического характера.

Технология использования критерия «восходящих» и «нисходящих» серий:

1 шаг. Сравнить каждое значение y_t с последующим значением y_{t+1} и сформировать последовательность S_t по правилу:

$$S_t = \begin{cases} <+> \quad при\ y_t < y_{t+1} \\ <-> \quad при\ y_t > y_{t+1} \\ не\ включается \quad при\ y_t = y_{t+1} \end{cases}, \qquad t = \overline{1, n-1}.$$

Результаты удобнее представить в виде таблицы:

t	y_t	S_t
1		
…		
n-1		

2 шаг. Выделить серии. Серией называется последовательность подряд идущих знаков «+» (восходящая серия) или «–» (нисходящая серия). Подсчитать K_{max} (протяженность самой длинной серии) и V (общее число серий);

3 шаг. Проверить выполнение неравенств:

– если следующие неравенства верны

$$K_{\max} < K_0,$$

$$V > \left[\frac{1}{3}\left(2n - 1 - 1{,}96\sqrt{\frac{16n-29}{90}}\right)\right], \text{ где [] – целая часть числа,}$$

то с 5%-ым уровнем значимости гипотеза о наличии тренда принимается. K_0 определяется в зависимости от n:

n	$n \le 26$	$26 < n \le 153$	$153 < n \le 1170$
K_0	5	6	7

– если хотя бы одно из этих неравенств нарушается, то гипотеза о наличии тенденции во временном ряду y_t отвергается.

Замечание 3. Количество серий не должно быть мало, а максимальная длина серии – велика. Если оба неравенства выполнены – делается вывод о существовании тренда, при нарушении хотя бы одного – о невозможности определения наличия тренда данным методом.

Пример 4. Имеется временной ряд y_t, $t = \overline{1, 49}$, содержащий данные об объеме финансирования в отрасль связи, млн. руб.(таб. 13).

Таблица 13

Исходные данные примера 4

№	y_t	№	y_t
1	22	16	45
2	24	17	50
3	26	18	50
4	26	19	50
5	32	20	51
6	27	21	50
7	33	22	56
8	35	23	54
9	39	24	57
10	37	25	57
11	41	26	59
12	43	27	55
13	41	28	61
14	47	29	61
15	49	30	62
		31	61

Определить наличие тренда в y_t *с помощью* критерия «восходящих» и «нисходящих» серий.

1 шаг. Составить последовательность S_t *(*таб. *14).*

Таблица 14

Пример составления последовательности S_t для критерия «восходящих» и «нисходящих» серий

№	y_t	S_t
1	22	
2	24	+
3	26	+
4	26	+
5	32	+
6	27	-
7	33	+
8	35	+
9	39	+
10	37	-
11	41	+
12	43	+
13	41	-
14	47	+
15	49	+

№	y_t	S_t
16	45	-
17	50	+
18	50	
19	50	
20	51	+
21	50	-
22	56	+
23	54	-
24	57	+
25	57	
26	59	+
27	55	-
28	61	+
29	61	
30	62	+
31	61	-
32	65	+
33	62	-
34	68	+
35	69	+
36	70	+
37	70	
38	71	+
39	71	
40	71	
41	78	+
42	80	+
43	82	+
44	91	+
45	93	+
46	97	+
47	100	+
48	98	-
49	97	-

2 шаг. Рассчитать количество серий V и максимальную длину K_{max} *(таб. 15).*
V= 20
$K_{max}=11$

Таблица 15

Пример расчета количества серий и длины серии в примере 4

№	y_t	S_t	Количество серий
1	22		
2	24	+	1
3	26	+	
4	26	+	
5	32	+	
6	27	-	2
7	33	+	3
8	35	+	
9	39	+	
10	37	-	4
11	41	+	5
12	43	+	
13	41	-	6
14	47	+	7
15	49	+	
16	45	-	8
17	50	+	9
18	50		
19	50		
20	51	+	
21	50	-	10
22	56	+	11
23	54	-	12
24	57	+	13
25	57		
26	59	+	
27	55	-	14
28	61	+	15
29	61		
30	62	+	
31	61	-	16
32	65	+	17
33	62	-	18
34	68	+	19
35	69	+	
36	70	+	
37	70		
38	71	+	
39	71		
40	71		

№	y_t	S_t	Количество серий
1	22		
41	78	+	
42	80	+	
43	82	+	
44	91	+	
45	93	+	
46	97	+	
47	100	+	
48	98	-	20
49	97	-	
K_{max}			11

Шаг 3. Проверить выполнение неравенств.

Рассчитать значение, с которым сравнивается V (рис. 5). Так как в нашем случае $K_0=6< K_{max}=11$ *и* $V=20> 14$, *то ни одно из вышеприведенных неравенств не выполняется, и, следовательно,* гипотеза о наличии тенденции во временном ряду y_t не подтверждается.

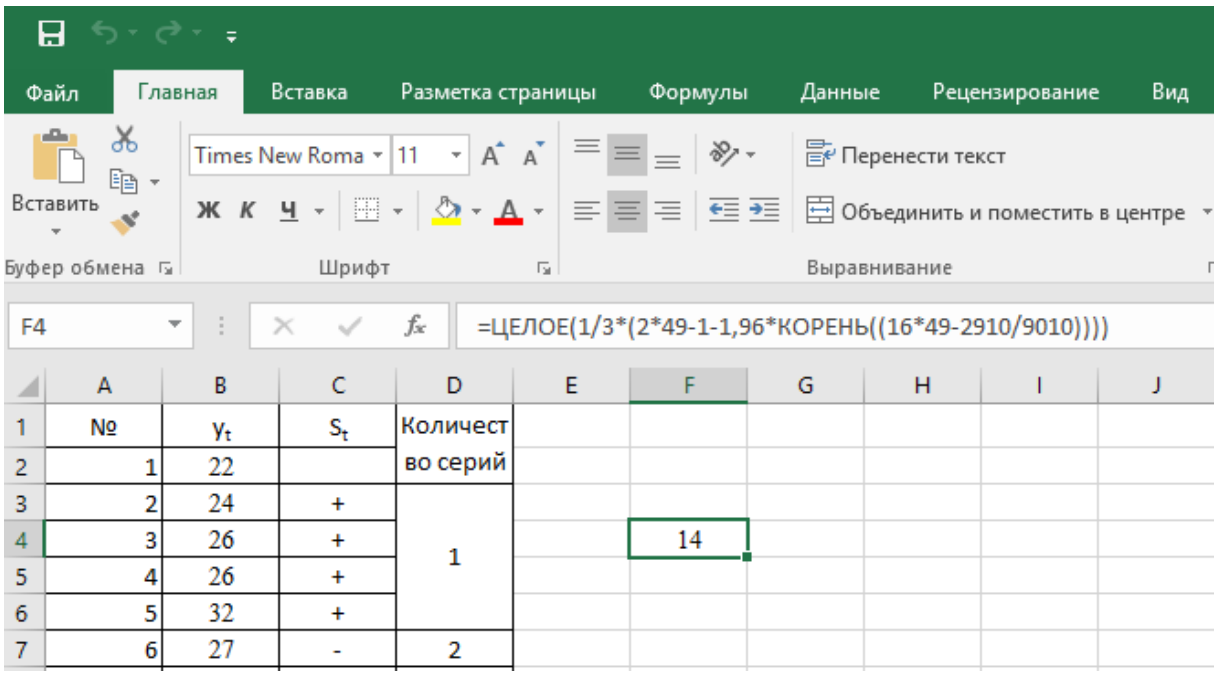


Рис. 5. Пример вычисления расчетного значения для критерия «восходящих» и «нисходящих» серий

Критерий серий

Критерий серий используется для проверки утверждения о присутствии во временном ряду трендовой компоненты, он основывается на медиане выборочной совокупности. Временной ряд объёмом *n* ранжируется, т. е. все наблюдения

упорядочиваются по возрастанию, и рассчитывается медиана ранжированного ряда.

Медианой называется наблюдение, которое делит ранжированный временной ряд на две равные части.

Если временной ряд содержит нечётное количество наблюдений, то в качестве медианы принимается значение, стоящее в середине данного ряда.

Если временной ряд содержит чётное количество наблюдений, то в качестве медианы берётся среднее арифметическое значение двух наблюдений, находящихся посередине временного ряда.

Технология использования критерия серий:

1 шаг. Уровни временного ряда y_t, $t=\overline{1,n}$ расположить по возрастанию и найти выборочную медиану y_{med} ряда:

$$y_{med}=\begin{cases} y_{\frac{n+1}{2}}, & \text{если } n \text{ нечетное}, \\ \frac{1}{2}\left(y_{\frac{n}{2}}+y_{\frac{n}{2}+1}\right), & \text{если } n \text{ четное}. \end{cases}$$

Замечание 4. Если для подсчёта медианы используется встроенная функция =МЕДИАНА(), то сортировка уровней временного ряда не нужна.

2 шаг. Сравнить каждое значение ряда y_t с y_{med} и сформировать последовательность S_t, состоящую из знаков «+» и «-» по правилу:

$$S_t=\begin{cases} <+> & \text{при } y_t > y_{med} \\ <-> & \text{при } y_t < y_{med} \\ \text{не включается} & \text{при } y_t = y_{med} \end{cases}, \qquad t=\overline{1,n}$$

Результаты удобнее представить в виде таблицы:

t	y_t	S_t
1		
…		
n		

3 шаг. Выделить серии. Серией называется последовательность подряд идущих знаков «+» или «–». Подсчитать K_{max} (протяженность самой длинной серии) и V (общее число серий);

4 шаг. Проверить выполнение гипотезы о наличии тренда:

– если следующие неравенства верны

$$K_{\max} < [1{,}43\ln(n+1)],$$

$$V > \left[\frac{1}{2}\left(n+2-1{,}96\sqrt{n-1}\right)\right],$$

где [] – целая часть числа, то с 5%-ым уровнем значимости гипотеза принимается;

– если хотя бы одно из этих неравенств нарушается, то гипотеза о наличии тенденции во временном ряду y_t отвергается.

Пример 5. Имеется временной ряд y_t, $t = \overline{1, 46}$, содержащий данные об удельном весе автомобильного транспорта в объеме перевозок грузов транспортом общего пользования, % (таб. 16).

Таблица 16

Исходные данные для примера 5

№	y_t	№	y_t
1	10	24	46
2	13	25	39
3	19	26	37
4	24	27	43
5	22	28	50
6	22	29	40
7	26	30	48
8	23	31	48
9	19	32	49
10	17	33	46
11	20	34	47
12	22	35	38
13	18	36	47
14	23	37	48
15	29	38	54
16	27	39	56
17	29	40	56
18	25	41	56
19	31	42	57
20	33	43	59
21	35	44	54
22	46	45	55
23	32	46	57

Требуется определить наличие тренда с помощью критерия серий.

1 шаг. Рассчитать медиану с помощью стандартной функции =МЕДИАНА (в качестве аргумента указать диапазон с уровнями временного ряда), в этом случае упорядочивать временной ряд по возрастанию значений не нужно.

2 шаг. Составить последовательность S_t, состоящую из знаков «+» и «-» (таб. 17).

Таблица 17

Пример составления последовательности St для критерия серий

№	y_t	S_t
1	10	-
2	13	-
3	19	-
4	24	-
5	22	-
6	22	-
7	26	-
8	23	-
9	19	-
10	17	-
11	20	-
12	22	-
13	18	-
14	23	-
15	29	-
16	27	-
17	29	-
18	25	-
19	31	-
20	33	-
21	35	-
22	46	+
23	32	-
24	46	+
25	39	+
26	37	-
27	43	+
28	50	+
29	40	+
30	48	+
31	48	+
32	49	+
33	46	+
34	47	+
35	38	+
36	47	+
37	48	+
38	54	+
39	56	+
40	56	+

№	y_t	S_t
41	56	+
42	57	+
43	59	+
44	54	+
45	55	+
46	57	+
Медиана	37,5	

3 шаг. Определить количество V и максимальную протяженность самой длинной серий K_{max}. После расчетов ⇒ V=6, $K_{max}=21$ (таб. 18).

Таблица 18

Пример расчета количества серий и максимальной протяженности серии критерия серий

№	y_t	S_t	Количество серий
1	10	-	1
2	13	-	
3	19	-	
4	24	-	
5	22	-	
6	22	-	
7	26	-	
8	23	-	
9	19	-	
10	17	-	
11	20	-	
12	22	-	
13	18	-	
14	23	-	
15	29	-	
16	27	-	
17	29	-	
18	25	-	
19	31	-	
20	33	-	
21	35	-	
22	46	+	2
23	32	-	3
24	46	+	4
25	39	+	
26	37	-	5
27	43	+	6

№	y_t	S_t	Количество серий
28	50	+	
29	40	+	
30	48	+	
31	48	+	
32	49	+	
33	46	+	
34	47	+	
35	38	+	
36	47	+	
37	48	+	
38	54	+	
39	56	+	
40	56	+	
41	56	+	
42	57	+	
43	59	+	
44	54	+	
45	55	+	
46	57	+	
Ме-диана	37,5	K_{max}	21

4 шаг. Проверить выполнение неравенств, для этого рассчитать по формулам соответствующие значения,

=ЦЕЛОЕ(1,43*LN(46+1)) – для проверки длины серии,
=ЦЕЛОЕ(1/2*(46+2-1,96*КОРЕНЬ(46-1))) – для проверки количества серий
Сравнить полученные значения:

K_{max}	>	5
V	<	17

Ни одно из проверочных неравенств не выполняется, следовательно, гипотеза о наличии тенденции во временном ряду y_t не подтверждается.

Методы механического сглаживания

Сглаживание временного ряда, т. е. замена фактических уровней расчетными значениями, имеющими меньшую колеблемость, чем исходные данные является простым методом выявления тенденции развития. Соответствующее преобразование называется фильтрованием.

Сглаживание временных рядов проводится по многим причинам. В ряде случаев при графическом изображении временного ряда тренд прослеживается недостаточно отчетливо. Поэтому ряд сглаживают, на график наносят сглаженные значения и, как правило, тенденция проявляется более четко. Некоторые методы анализа и прогнозирования требуют в качестве предварительного условия сглаживание временного ряда. Сглаживание временных рядов используется при устранении аномальных наблюдений. Методы сглаживания в настоящее время применяются для непосредственного прогнозирования экономических показателей.

Исключение случайных колебаний значений уровней ряда осуществляется с помощью нахождения «усредненных» значений. Способы устранения случайных факторов делятся на две больше группы:

1. Способы «механического» сглаживания колебаний путем усреднения значений ряда относительно других, расположенных рядом, уровней ряда.

2. Способы «аналитического» выравнивания, т. е. определения сначала функционального выражения тенденции ряда, а затем новых, расчетных значений ряда.

Метод простой скользящей средней

Данный метод применяется для характеристики тенденции развития исследуемой статистической совокупности и основан на расчете средних уровней ряда за определенный период. Согласно этому методу определяется количество наблюдений m, входящих в интервал сглаживания. При этом используют правило: если необходимо сгладить мелкие, беспорядочные колебания, то интервал сглаживания берут по возможности большим и, наоборот, интервал сглаживания уменьшают, когда нужно сохранить более мелкие волны и освободиться от периодически повторяющихся колебаний, возникающих, например, из-за автокорреляций уровней.

В результате использования метода простой скользящей средней получается $n\text{-}m+1$ сглаженных значений уровней ряда; при этом первые p и последние p уровней ряда теряются (не сглаживаются).

Для вычисления сглаженных уровней ряда при нечетном значении интервала сглаживания m применяется формула:

$$\hat{y}_t = \frac{\sum_{i=t-p}^{t+p} y_i}{m}, \quad t > p, \quad p = \frac{m-1}{2}.$$

В случае четного значения интервала сглаживание осуществляется по формуле:

$$\hat{y}_t = \frac{\frac{1}{2} y_{t-p} + \sum_{i=t-p+1}^{t+p-1} y_i + \frac{1}{2} y_{t+p}}{m}, \quad t > p, \quad p = \frac{m}{2}.$$

Замечание 5. Прогноз на основе скользящего среднего имеет ту особенность, что всем данным, включенным в процесс скользящего среднего, приписывается одинаковый вес. Всем остальным данным приписывается нулевой вес.

Вес отдельного наблюдения характеризует долю его вклада при расчете скользящего среднего. При этом более старые данные имеют тот же вес, что и более свежие, хотя понятно, что свежие данные имеют более важное значение с точки зрения формулирования прогнозных оценок.

Замечание 6. Метод простой скользящей средней возможно использовать, если графическое изображение ряда напоминает прямую линию. В этом случае не искажается динамика развития исследуемого процесса. Однако, когда тренд выравниваемого ряда имеет изгибы и к тому же желательно сохранить мелкие волны, использовать для сглаживания ряда метод простой скользящей средней нецелесообразно, поскольку при этом выравниваются и выпуклые, и вогнутые линии; происходит сдвиг волны вдоль ряда; изменяется знак волны, т. е. на кривой, соединяющей сглаженные точки, вместо выпуклого участка образуется вогнутый и наоборот. Последнее имеет место в случаях, когда интервал сглаживания в полтора раза превышает длину волны.

Таким образом, если развитие процесса носит нелинейный характер, то применение метода простой скользящей средней может привести к значительным искажениям исследуемого процесса. В таких случаях более надежным является использование других методов сглаживания, например, метода взвешенной скользящей средней.

Пример 6. Имеется временной ряд y_t, $t = \overline{1, 46}$, содержащий данные об удельном весе автомобильного транспорта в объеме перевозок грузов транспортом общего пользования, % (таб. 19).

Таблица 19

Исходные данные для примера 6

№	y_t	№	y_t	№	y_t
1	10	17	29	33	46
2	13	18	25	34	47
3	19	19	31	35	38
4	24	20	33	36	47
5	22	21	35	37	48
6	22	22	46	38	54
7	26	23	32	39	56
8	23	24	46	40	56
9	19	25	39	41	56
10	17	26	37	42	57
11	20	27	43	43	59
12	22	28	50	44	54
13	18	29	40	45	55
14	23	30	48	46	57
15	29	31	48		
16	27	32	49		

Требуется сгладить временной ряд с помощью простой скользящей средней (период сглаживания $m = 3$).

Рассчитаем количество p уровней, которые теряются (не сглаживаются) для периода сглаживания $m = 3$, получим $p = 1$, следовательно, начиная со второго уровня и без последнего элементы ряда пересчитываются (таб. 20):

Таблица 20

Пример сглаживания уровней временного ряда методом скользящей средней

№	y_t	$y_{t\ сглаж.}$
1	10	
2	13	14
3	19	18,66667
4	24	21,66667
5	22	22,66667
6	22	23,33333
7	26	23,66667
8	23	22,66667
9	19	19,66667
10	17	18,66667
11	20	19,66667
12	22	20
13	18	21
14	23	23,33333
15	29	26,33333

№	y_t	$y_{t\ сглаж.}$
16	27	28,33333
17	29	27
18	25	28,33333
19	31	29,66667
20	33	33
21	35	38
22	46	37,66667
23	32	41,33333
24	46	39
25	39	40,66667
26	37	39,66667
27	43	43,33333
28	50	44,33333
29	40	46
30	48	45,33333
31	48	48,33333
32	49	47,66667
33	46	47,33333
34	47	43,66667
35	38	44
36	47	44,33333
37	48	49,66667
38	54	52,66667
39	56	55,33333
40	56	56
41	56	56,33333
42	57	57,33333
43	59	56,66667
44	54	56
45	55	55,33333
46	57	

Визуальное сравнение исходных и выровненных уровней временного ряда приведено на рис. 6. Мелкие, беспорядочные и периодически повторяющиеся колебания сглажены.

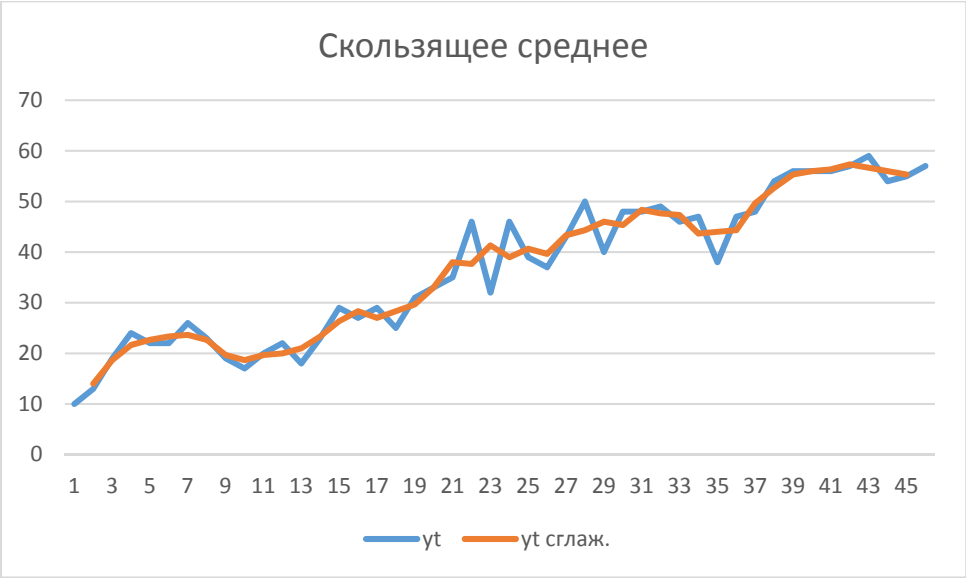


Рис. 6. Диаграмма сравнения исходных и сглаженных уровней временного ряда

Рассмотрим использование пакета Анализ данных в табличном процессоре Excel для выравнивания уровней временного ряда методом скользящей средней. Выбор команды **Данные/Анализ данных/Скользящая средняя** выводит диалоговое окно (рис 7), в котором задаются входные данные для расчета скользящего среднего:

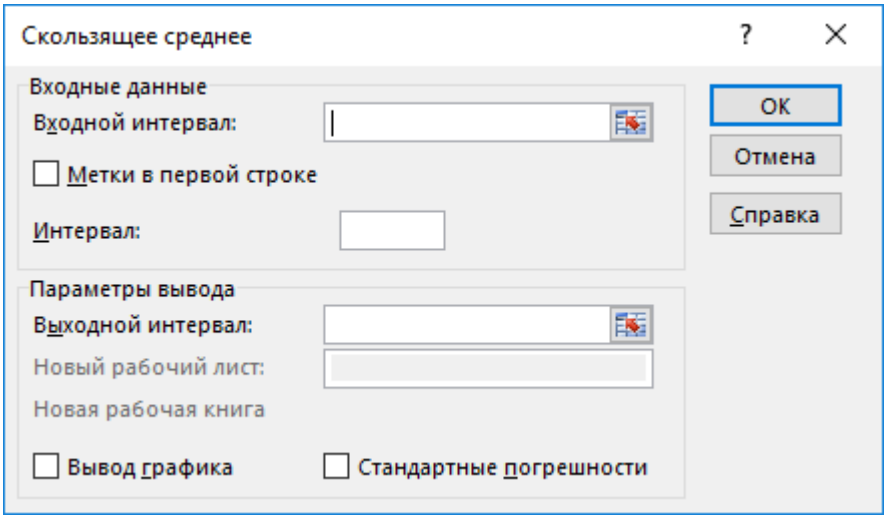


Рис. 7. Диалоговое окно Скользящая средняя

– входной интервал – диапазон ячеек, в котором заданы уровни временного ряда;

– метки – указать в том случае, если входные данные начинаются с имени ряда;

– интервал – период сглаживания;

– выходной интервал – ячейка, с которой начинается вывод сглаженных значений;

– вывод графика – показать диаграмму для визуального сравнения исходных и выровненных уровней временного ряда;

– стандартные погрешности –стандартные погрешности сглаженных уровней ряда.

После применения команды выводится диапазон ячеек сглаженных уровней ряда (табл. 21) и диаграмма для визуального сравнения исходных и выровненных уровней временного ряда (рис. 8), если задан вывод графика.

Таблица 21

Пример вывода сглаженных уровней временного ряда методом скользящей средней

#Н/Д
#Н/Д
14
18,66667
21,66667
22,66667
23,33333
23,66667
22,66667
19,66667
18,66667
19,66667
20
21
23,33333
26,33333
28,33333
27
28,33333
29,66667
33
38
37,66667
41,33333
39
40,66667
39,66667
43,33333
44,33333
46
45,33333
48,33333

47,66667
47,33333
43,66667
44
44,33333
49,66667
52,66667
55,33333
56
56,33333
57,33333
56,66667
56
55,33333

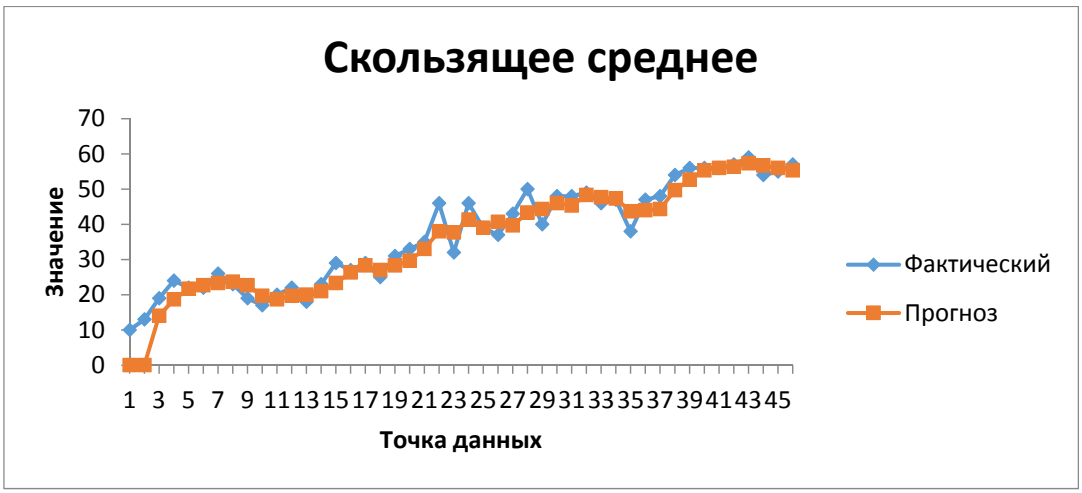


Рис. 8. Диаграмма визуального сравнения исходных и выровненных уровней временного ряда

Метод взвешенной скользящей средней

Этот метод отличается от предыдущего тем, что сглаживание внутри интервала производится не по прямой, а по кривой более высокого порядка. Это обусловлено тем, что суммирование членов ряда, входящих в интервал сглаживания, производится с определенными весами, рассчитанными по методу наименьших квадратов.

$$\tilde{y}_t = \frac{\sum_{t=t-p}^{t+p} \omega_t y_t}{\sum_{t=t-p}^{t+p} \omega_t},$$

Значения весовых коэффициентов для различных длин отрезков *m* и порядка полинома приведены в таблице 22.

Таблица 22

Значение некоторых весовых коэффициентов

m	p	Значения весовых коэффициентов
m_0	0 или 1	$\frac{1}{2m_0+1}$
5	2 или 3	$-\frac{3}{35};\ \frac{17}{35};\ \frac{12}{35};\ \frac{12}{35};\ -\frac{3}{35};$
7	2 или 3	$-\frac{2}{21};\ \frac{3}{21};\ \frac{6}{21};\ \frac{7}{21};\ \frac{6}{21};\ \frac{3}{21};\ -\frac{2}{21}$
9	2 или 3	$-\frac{21}{231};\ \frac{14}{231};\ \frac{39}{231};\ \frac{54}{231};\ \frac{59}{231};\ \frac{54}{231};\ \frac{39}{231};\ \frac{14}{231};\ -\frac{21}{231}$
7	4 или 5	$\frac{5}{231};\ -\frac{30}{231};\ \frac{75}{231};\ \frac{131}{231};\ \frac{75}{231};\ -\frac{30}{231};\ \frac{5}{231}$
9	4 или 5	$\frac{15}{429};\ -\frac{55}{429};\ \frac{30}{429};\ \frac{135}{429};\ \frac{179}{429};\ \frac{135}{429};\ \frac{30}{429};\ -\frac{55}{429};\ \frac{15}{429}$

Метод экспоненциального сглаживания

Рассмотренные методы простой и взвешенной скользящей средней не дают возможности сгладить первые и последние *p* наблюдений временного ряда. Отсутствие сглаженных первых наблюдений не так важно по сравнению с последними наблюдениями, особенно если целью исследования является прогнозирование развития процесса. Есть методы, позволяющие получить сглаженные значения последних уровней так же, как и всех остальных. К их числу относится метод экспоненциального сглаживания. Экспоненциальное сглаживание является одним из наиболее распространенных приемов, используемых для сглаживания временных рядов, а также для прогнозирования. В основе процедуры сглаживания лежит расчёт экспоненциальных скользящих средних сглаживаемого ряда.

Главное достоинство прогнозной модели, основанной на экспоненциальных средних, состоит в том, что она способна последовательно адаптироваться к новому уровню процесса без значительного реагирования на случайные отклонения.

В методе простого экспоненциального сглаживания применяется взвешенное (экспоненциально) скользящее усреднение всех данных предыдущих наблюдений. Эта модель чаще всего применяется к данным, в которых необходимо оценить наличие зависимости между анализируемыми показателями (тренда) или зависимость анализируемых данных. Целью экспоненциального сглаживания является оценка текущего состояния, результаты которого определят все последующие прогнозы. Экспоненциальное сглаживание предусматривает постоянное обновление модели за счет наиболее свежих данных. Этот метод основывается на усреднении (сглаживании) временных рядов прошлых наблюдений в нисходящем (экспоненциально) направлении. То есть более поздним событиям

присваивается больший вес. Вес присваивается следующим образом $\omega_t = \gamma(1-\gamma)^{n-t}$.

Процедура простого экспоненциального сглаживания осуществляется по следующим формулам: $\hat{y}_1 = y_1$, $\hat{y}_t = \gamma \cdot \hat{y}_{t-1} + (1-\gamma) \cdot y_t$, $t = \overline{2,n}$, где:

y_t – фактическое наблюдение в момент t;

$\hat{y}_t$ – значение экспоненциального среднего в момент t;

γ – параметр сглаживания, $\gamma = const$, $\gamma \in (0; 1]$ задается априорно или может быть рассчитан по формуле $\gamma = \frac{2}{m+1}$, m – период сглаживания. Величина $(1-\gamma)$ называется коэффициентом дисконтирования.

Экспоненциальное среднее в момент t здесь выражено как взвешенная сумма текущего наблюдения и экспоненциального среднего прошлого наблюдения с весами γ и $(1 - \gamma)$ соответственно. Если последовательно использовать данное рекуррентное соотношение, то значение $\hat{y}_t$можно выразить через значения временного ряда y_t: $\hat{y}_t = \gamma \sum_{i=0}^{\infty} (1-\gamma)^i y_{t-1}$

Таким образом, величина $\hat{y}_t$оказывается взвешенной суммой всех членов ряда. Причем значения весов уменьшаются экспоненциально в зависимости от удаленности наблюдения относительно момента t. Это и объясняет название «экспоненциальное среднее».

Экспоненциальное сглаживание можно представить как фильтр, на вход которого в виде потока последовательно поступают члены исходного ряда, а на выходе формируются значения экспоненциальных средних. Причем, сглаженный ряд $\hat{y}_t$имеет тоже математическое ожидание, что и ряд y_t, но меньшую дисперсию.

При высоком значении γ дисперсия сглаженного ряда не значительно отличается от дисперсии ряда y_t. Чем меньше γ, тем в большей степени сокращается дисперсия сглаженного ряда (то есть подавляются колебания исходного ряда).

При вычислении $\hat{y}_1$ начальное значение параметра $\hat{y}_0$ принимается либо равным значению уровня ряда y_1, либо равным средней арифметической нескольких первых членов ряда, например, y_1, y_2, y_3:

$$\hat{y}_0 = \frac{y_1 + y_2 + y_3}{3}.$$

Пример 7. Имеется временной ряд y_t, $t = \overline{1,46}$, содержащий данные об удельном весе автомобильного транспорта в объеме перевозок грузов транспортом общего пользования, %. Требуется сгладить временной ряд с помощью экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}7$) (таб. 23).

Таблица 23

Исходные данные примера 7

№	y_t	№	y_t	№	y_t
1	10	17	29	33	46
2	13	18	25	34	47
3	19	19	31	35	38
4	24	20	33	36	47

№	y_t	№	y_t	№	y_t
5	22	21	35	37	48
6	22	22	46	38	54
7	26	23	32	39	56
8	23	24	46	40	56
9	19	25	39	41	56
10	17	26	37	42	57
11	20	27	43	43	59
12	22	28	50	44	54
13	18	29	40	45	55
14	23	30	48	46	57
15	29	31	48		
16	27	32	49		

1 шаг. Рассчитать выровненные уровни временного ряда для параметра сглаживания γ=0,7 (таб. 24).

Таблица 24

Пример сглаживания уровней временного ряда экспоненциальным методом

№	y_t	$y_{t\ экспон}$
1	10	10
2	13	10,9
3	19	13,33
4	24	16,531
5	22	18,1717
6	22	19,32019
7	26	21,32413
8	23	21,82689
9	19	20,97883
10	17	19,78518
11	20	19,84962
12	22	20,49474
13	18	19,74632
14	23	20,72242
15	29	23,20569
16	27	24,34399
17	29	25,74079
18	25	25,51855
19	31	27,16299
20	33	28,91409
21	35	30,73986
22	46	35,3179
23	32	34,32253
24	46	37,82577
25	39	38,17804
26	37	37,82463

№	y_t	$y_{t\ экспон}$
27	43	39,37724
28	50	42,56407
29	40	41,79485
30	48	43,65639
31	48	44,95948
32	49	46,17163
33	46	46,12014
34	47	46,3841
35	38	43,86887
36	47	44,80821
37	48	45,76575
38	54	48,23602
39	56	50,56522
40	56	52,19565
41	56	53,33696
42	57	54,43587
43	59	55,80511
44	54	55,26358
45	55	55,1845
46	57	55,72915

Визуальное сравнение исходных и выровненных уровней временного ряда приведено на рис. 9. Мелкие, беспорядочные и периодически повторяющиеся колебания сглажены.

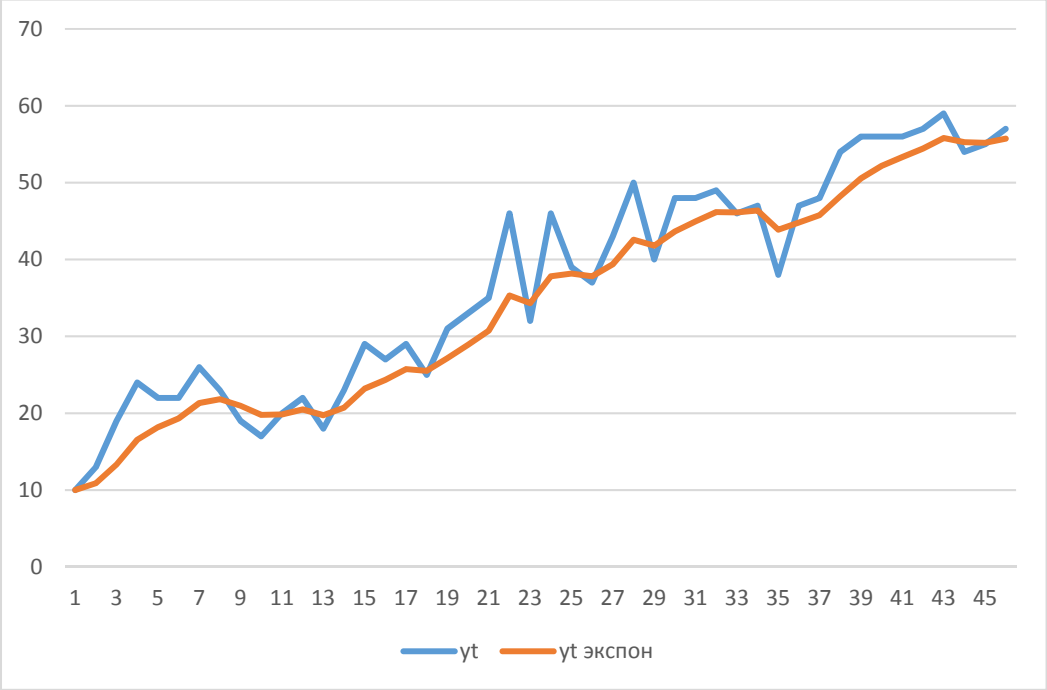


Рис. 9. Диаграмма визуального сравнения исходных и экспоненциально выровненных уровней временного ряда

Расчет весовых коэффициентов приведен в таблице 25. Причем значения весов уменьшаются экспоненциально в зависимости от удаленности наблюдения относительно момента *t* (рис. 10)

Таблица 25

Расчет весовых коэффициентов

№	y_t	$y_{t\ экспон}$	Весовые коэф-ты
1	10	10	2,06802E-24
2	13	10,9	6,8934E-24
3	19	13,33	2,2978E-23
4	24	16,531	7,65933E-23
5	22	18,1717	2,55311E-22
6	22	19,32019	8,51037E-22
7	26	21,32413	2,83679E-21
8	23	21,82689	9,45596E-21
9	19	20,97883	3,15199E-20
10	17	19,78518	1,05066E-19
11	20	19,84962	3,50221E-19
12	22	20,49474	1,1674E-18
13	18	19,74632	3,89134E-18
14	23	20,72242	1,29711E-17
15	29	23,20569	4,32371E-17
16	27	24,34399	1,44124E-16
17	29	25,74079	4,80413E-16
18	25	25,51855	1,60138E-15
19	31	27,16299	5,33792E-15
20	33	28,91409	1,77931E-14
21	35	30,73986	5,93102E-14
22	46	35,3179	1,97701E-13
23	32	34,32253	6,59002E-13
24	46	37,82577	2,19667E-12
25	39	38,17804	7,32225E-12
26	37	37,82463	2,44075E-11
27	43	39,37724	8,13583E-11
28	50	42,56407	2,71194E-10
29	40	41,79485	9,03981E-10
30	48	43,65639	3,01327E-09
31	48	44,95948	1,00442E-08
32	49	46,17163	3,34808E-08
33	46	46,12014	1,11603E-07
34	47	46,3841	3,72009E-07
35	38	43,86887	1,24003E-06
36	47	44,80821	4,13343E-06
37	48	45,76575	1,37781E-05
38	54	48,23602	4,5927E-05
39	56	50,56522	0,00015309

№	y_t	$y_{t\ экспон}$	Весовые коэф-ты
40	56	52,19565	0,0005103
41	56	53,33696	0,001701
42	57	54,43587	0,00567
43	59	55,80511	0,0189
44	54	55,26358	0,063
45	55	55,1845	0,21
46	57	55,72915	0,7

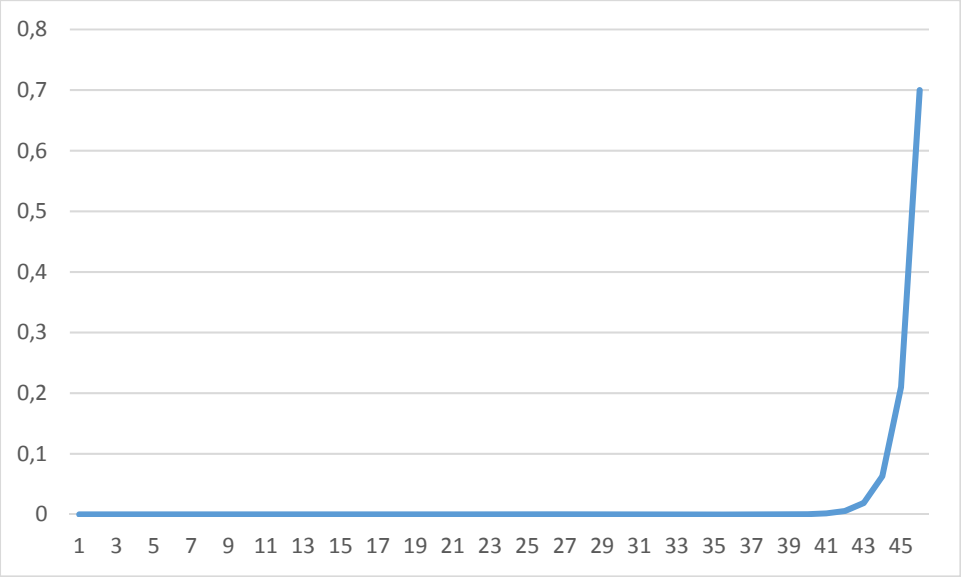


Рис. 10. Графическое представление весовых коэффициентов

Приведем процедуру экспоненциального сглаживания с помощью табличного процессора Excel. Требуется сгладить временной ряд примера 7 с помощью экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}7$). Используем команду **Данные/Анализ данных/Экспоненциальное сглаживание**, в диалоговом окне (рис 11) задаются входные данные для расчета уровней ряда:

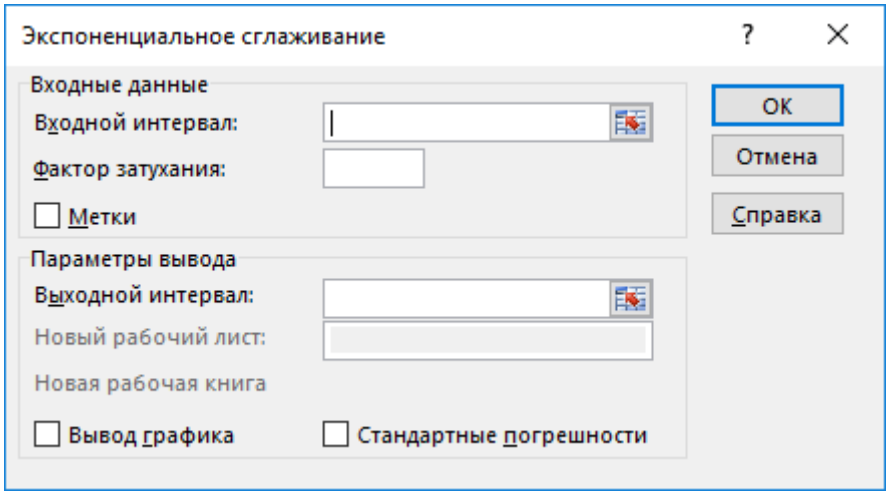


Рис. 11. Диалоговое окно Экспоненциальное сглаживание

–входной интервал – диапазон ячеек, в котором заданы уровни временного ряда;

– Метки – указть в том случае, если входные данные начинаются с имени ряда;

– Фактор затухания – параметр сглаживания;

– Выходной интервал – ячейка, с которой начинается вывод сглаженных значений;

– Вывод графика – показать диаграмму для визуального сравнения исходных и выровненных уровней временного ряда;

– Стандартные погрешности –стандартные погрешности сглаженных уровней ряда.

После применения команды выводится диапазон ячеек сглаженных уровней ряда (таб. 26) и диаграмма для визуального сравнения исходных и выровненных уровней временного ряда (рис. 13), если задан вывод графика.

Таблица 26

Выровненные уровни временного ряда метода экспоненциального сглаживания

№	y_t	$y_{t\ экспон}$
1	10	10
2	13	10,9
3	19	13,33
4	24	16,531
5	22	18,1717
6	22	19,32019
7	26	21,32413
8	23	21,82689
9	19	20,97883
10	17	19,78518
11	20	19,84962
12	22	20,49474
13	18	19,74632
14	23	20,72242
15	29	23,20569
16	27	24,34399
17	29	25,74079
18	25	25,51855
19	31	27,16299
20	33	28,91409
21	35	30,73986
22	46	35,3179
23	32	34,32253
24	46	37,82577
25	39	38,17804
26	37	37,82463
27	43	39,37724
28	50	42,56407

№	y_t	$y_{t\ экспон}$
29	40	41,79485
30	48	43,65639
31	48	44,95948
32	49	46,17163
33	46	46,12014
34	47	46,3841
35	38	43,86887
36	47	44,80821
37	48	45,76575
38	54	48,23602
39	56	50,56522
40	56	52,19565
41	56	53,33696
42	57	54,43587
43	59	55,80511
44	54	55,26358
45	55	55,1845
46	57	55,72915

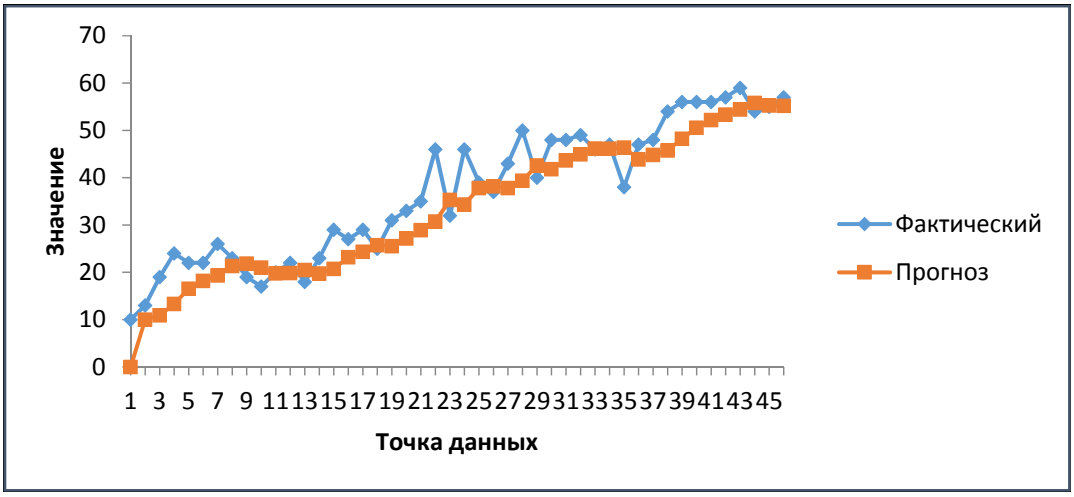


Рис. 12. Сравнение выровненных и исходных уровней временного ряда метода экспоненциального сглаживания

Обоснование выбора уравнения тренда

Более точным способом отображения тенденции динамического ряда является аналитическое выравнивание, т. е. выравнивание с помощью аналитических формул. В этом случае динамический ряд выражается в виде функции $y=f(t)$, в которой в качестве основного фактора принимается время t, и изменения аргумента функции определяют расчетные значения y_t.

Фактическими (или эмпирическими) уровнями ряда динамики называют исходные данные об изменении явления, т. е. данные, полученные опытным путем, посредством наблюдения. Они обозначаются y_t. Расчетными (или теоретическими) уровнями ряда называют значения, полученные в результате подстановки в уравнение тренда значений t, и обозначают их $\hat{y}_t$.

Целью аналитического выравнивания динамического ряда является определение аналитической или графической зависимости $f(t)$. На практике по имеющемуся временному ряду задают вид и находят параметры функции $f(t)$, а затем анализируют поведение отклонений от тенденции. Функцию $f(t)$ выбирают таким образом, чтобы она давала содержательное объяснение изучаемого процесса.

Экстраполяция тенденций динамических рядов широко применяется на практике в силу ее простоты и возможности осуществления на основе относительно небольшого объема информации. Она допустима и правомерна при выполнении следующих условий:

1. Период предыстории, за который построена модель тренда, должен быть достаточным для выявления тенденции развития показателя;
2. Анализируемый процесс является устойчивым и обладает инерционностью, то есть для значительных изменений характеристик ряда (в частности, для перехода к другому типу развития) требуется продолжительный период;
3. Не ожидается сильных внешних воздействий на изучаемый процесс, которые могут серьезно повлиять на тенденцию развития.

Экстраполяция по трендам может применяться и как начальный этап комплексной методики прогнозирования, отвечающий на вопрос о последствиях продолжения прежней тенденции развития. В этом случае экстраполяционный прогноз интерпретируется как один из гипотетических вариантов, с которым сопоставляются другие варианты прогноза, полученные с помощью более совершенных методов.

Операцию экстраполяции по модели тренда можно представить так:

$$\hat{y}(n+\ell)=f(y_n,\ell,a_j),$$

где $\hat{y}(n+\ell)$ - прогнозируемый уровень ряда,

f - функция, с помощью которой описывают тренд,

y_n - уровень ряда, принятый за базу экстраполяции,

ℓ - период упреждения прогноза,

a_j - параметр уравнения тренда.

Процесс прогнозирования по модели тренда состоит из следующих этапов:

1. Выбор функции для описания тренда,

2. Оценка параметров модели (моделей) тренда,
3. Проверка адекватности и точности модели тренда,
4. Расчет точечного и интервального прогнозов.

Таким образом, целью аналитического выравнивания является:

– определение вида функционального уравнения;
– нахождения параметров уравнения;
– расчет «теоретических», выровненных уровней, отображающих основную тенденцию ряда динамики.

Графическое отображение изменения уровней ряда играет большую роль в применении данного вида выравнивания. Оно позволяет ускорить процедуру анализа и увеличить степень наглядности полученных результатов.

Выбор функции может быть осуществлен несколькими способами: логический анализ явления (процесса); визуальный анализ уровней ряда; формальный анализ ретроспективного ряда динамики.

Содержание *логического анализа* сводится к тому, чтобы понять природу динамики изучаемого явления (процесса). При проведении этого этапа стараются получить ответы на следующие вопросы:

– Каковы темпы развития процесса? В каком направлении пойдет развитие процесса в будущем?
– Возможно ли изменение темпов в периоде упреждения прогноза?
– Следует ли ожидать в развитии процесса скачкообразного изменения?
– Могут ли возникнуть точки перегиба?
– Существуют ли пределы развития экономического явления?

При *визуальном анализе* строят график, характеризующий изменение показателя в периоде предыстории, и сопоставляют этот график с графиками выбранного класса функций.

Графическое изображение позволяет осуществить контроль достоверности статистических показателей, так как, представленные на графике, они более ярко показывают имеющиеся неточности, связанные либо с наличием ошибок наблюдения, либо с сущностью изучаемого явления. С помощью графического изображения возможны изучение закономерностей развития явления, установление существующих взаимосвязей. Простое сопоставление данных не всегда дает возможность уловить наличие причинных зависимостей, в то же время их графическое изображение способствует выявлению причинных связей, в особенности в случае установления первоначальных гипотез, подлежащих затем дальнейшей разработке. Графики также широко используются для изучения структуры явлений, их изменения во времени и размещения в пространстве. В них более выразительно проявляются сравнительные характеристики и отчетливо виды основные тенденции развития и взаимосвязи, присущие изучаемому явлению или процессу.

При *формальном анализа* выбор функции тренда может быть осуществлен в ходе анализа *цепных абсолютных приростов* уровней ряда (первых разностей $\Delta_t = y_t - y_{t-1}, \quad t = \overline{2,n}$), *абсолютных ускорений уровней ряда* (вторых разностей

$\Delta_\Delta = \Delta_t - \Delta_{t-1}, \quad t = \overline{2,n}$), *цепных коэффициентов роста* ($K_t = y_t / y_{t-1}, \quad t = \overline{1,n}$). Если примерно одинаковы Δ_t, ряд имеет линейный тренд; если же примерно постоянны Δ_Δ, то для описания тенденции следует выбрать параболу второго порядка; если примерно равны K_t, следует использовать экспоненциальную или степенную функцию.

Одним из методов определения порядка полинома является метод последовательных разностей.

Метод последовательных разностей

Последовательно для $k = 1, 2, \ldots$ вычислить разности $\Delta^k y_t$:

$$\Delta^1 y_t = y_t - y_{t-1}, \quad t = \overline{2,n},$$
$$\Delta^2 y_t = \Delta^1 y_t - \Delta^1 y_{t-1}, \quad t = \overline{3,n},$$
$$\ldots$$
$$\Delta^k y_t = \Delta^{k-1} y_t - \Delta^{k-1} y_{t-1}, \quad t = \overline{k+1,n}.$$

Поведение разностей анализируется в зависимости от их порядка k. Если, начиная с некоторого k, разности стабилизируются, оставаясь приблизительно на одном уровне при дальнейшем росте k, это значение k и будет давать порядок сглаживающего полинома, то есть p.

Оценка параметров уравнения тренда методом наименьших квадратов

Для оценки параметров модели $a_0, a_1, \ldots$ чаще всего используется метод наименьших квадратов (МНК).

Идея МНК заключается в том, чтобы определить вид функции $\hat{y}_t$ в наибольшей степени соответствующей эмпирическим данным. Такая кривая должна обеспечить наименьшее значение суммы

$$S = \sum_{t=1}^{n} (y_t - \hat{y}_t)^2 \to \min.$$

Таким образом, меняя вид теоретических кривых $\hat{y}_t$, подбирают наименьшее значение S.

При этом следует иметь в виду, что МНК применим лишь в случае линейной зависимости функции и параметров. В том случае, когда во временных рядах проявляются тренды, описываемые функциями, нелинейными по параметрам, их оценка проводится с помощью МНК после линерализующих преобразований или нелинейного МНК.

Наиболее часто при прогнозировании социально-экономических явлений, имеющих тенденцию, применяют три вида функций: полиномиальные, экспоненциальные, S – образные.

Рассмотрим типы функций, наиболее часто используемые для выравнивания экономических временных рядов:

a. полиномиальные,
b. экспоненциальные,
c. S – образные.

Полиномиальные функции различных порядков имеют следующий вид:

$$y_t = a_0 + \sum_{i=1}^{p} a_i t^i \quad ,$$

где p – порядок полинома.

Из полиномиальных функций в экономике чаще всего используют полиномы первой, второй и третьей степени:

$$y_t = a_0 + a_1 t \quad ,$$
$$y_t = a_0 + a_1 t + a_2 t^2 \quad ,$$
$$y_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3 .$$

Линейная зависимость выбирается в тех случаях, когда в исходном временном ряду наблюдаются более или менее постоянные абсолютные и цепные приросты, не проявляющие тенденции ни к увеличению, ни к снижению.

Параболическая зависимость используется, если абсолютные цепные приросты сами по себе обнаруживают некоторую тенденцию развития, но абсолютные цепные приросты абсолютных цепных приростов (разности второго порядка) никакой тенденции развития не проявляют.

Параметры указанных полиномов имеют экономическую интерпретацию. Параметр a_0 характеризует уровень ряда при $t = 0$, параметр a_1 – средний прирост, параметр a_2 – скорость изменения среднего прироста, параметр a_3 – изменение ускорения роста.

Если рассчитать первые приросты $u_t = \hat{y}_t - \hat{y}_{t-1}$ для полинома первой степени, то они будут одинаковы и равны параметру a_1, для полиномов второй степени первые приросты будут иметь линейный характер, а вторые приросты $u_t^{[2]} = u_t - u_{t-1}$ будут постоянны и равны $2a_2$. У полиномов третьего порядка по линейному закону изменяются вторые приросты, а постоянны третьи приросты $u_t^{[3]} = u_t^{[2]} - u_{t-1}^{[2]}$. Таким образом, уравнение прямой применяют для описания процессов, равномерно изменяющихся во времени, параболу второй степени - для описания процессов, имеющих равноускоренный рост или равноускоренное снижение уровней ряда, кубическую параболу – когда наблюдается разный характер в изменении ускорения роста.

Экспоненциальные зависимости применяются, если в исходном временном ряду наблюдается либо более или менее постоянный относительный рост (устойчивость цепных темпов роста, темпов прироста, коэффициентов роста) , либо ,

при отсутствии такого постоянства, -- устойчивость в изменении показателей относительного роста (цепных темпов роста цепных же темпов роста , цепных коэффициентов роста цепных же коэффициентов или темпов роста и т.д.)

Простая экспонента

$$y_t = a_0 e^{a_1 t}$$

преобразуется к линейному виду с помощью логарифмирования $\ln y_t = \ln a_0 + a_1 t$ и введения новой переменной $z_t = ln\ y_t$. В результате линеаризации нелинейная функция приведена к линейному по параметрам и переменным виду.

Степенная функция

$$y_t = a_0 t^{a_1}$$

преобразуется к линейному виду с помощью логарифмирования и введения новых переменных $\ln y_t = \ln a_0 + a_1 \ln t$, $z_t = ln\ y_t$, $w_t = ln\ t$.

Логарифмическая функция

$$y_t = a_0 + a_1 \ln t$$

преобразуется к линейному виду с помощью введения новой переменной $z_t = ln\ t$.

S-образные функции используют для описания двух последовательных лавинообразных процесса: один с ускорением развития, другой - с замедлением. Графики их функций напоминают латинскую букву S, отсюда и название данного класса функций. Эти функции находят применение для описания демографических процессов, в страховании, могут применяться для описания процессов распространения новшеств и изобретений.

Функция

$$y_t = e^{a_0 + a_1 / t}$$

приводится к линейному виду с помощью логарифмирования и введения новой переменной:

$\ln y_t = a_0 + a_1 \vartheta_t$, $\vartheta_t = \frac{1}{t}$. $z_t = ln\ y_t$, $z_t = a_0 + a_1 \vartheta_t$

Прогнозирование уровней временного ряда

Обозначим через $\hat{y}_{t+l}$ оценку прогноза на l шагов вперед.

Точечный прогноз – это прогноз, который называет единственное значение прогнозируемого показателя. Значение точечного прогноза получается путем подстановки в выбранное уравнение величины времени t, соответствующего периоду упреждения ($t = n + 1;\quad t = n + 2;$ *и и.д.*).

Интервальный прогноз позволяет рассчитать интервал значений, в котором с достаточной долей уверенности можно ожидать уровень прогнозируемой величины.

Один из методов расчета границ интервала:

нижняя (левая) граница $\hat{y}_{t+l} - t_{табл} \cdot \sigma$

верхняя (правая) граница $\hat{y}_{t+l} + t_{табл} \cdot \sigma$

где $t_{табл} \cdot \sigma$ - доверительный полуинтервал;

$t_{табл}$ принимается при уровне значимости 0,05, и степени свобода *n-2*;

σ – стандартное отклонение фактических уровней ряда от рассчитанных по модели.

Результаты расчета прогнозов рекомендуется представить в виде таблицы:

Расчет прогнозов

Период T+l	***Точечный Прогноз***	***Нижняя граница***	***Верхняя Граница***
n + 1			
n + 2			
n + 3			
n + 4			

Оценка качества полученных результатов

Для сравнения различных альтернативных прогнозов необходим критерий оценки качества прогноза. Используются следующие критерии.

a. Коэффициент несовпадения ретроспективного предсказания $\hat{y}_t$ с наблюдавшимися значениями y_t, предложенный Тейлом:

$$L = \frac{\sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}\left(y_t - \hat{y}_t\right)^2\right]}}{\sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}\hat{y}_t^2\right]} + \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}y_t^2\right]}}$$

Значения коэффициента принадлежат отрезку [0, 1], причем на концах отрезка он имеет следующую содержательную интерпретацию: при L = 0 - отличное качество прогноза; при L = 1 – плохое качество прогноза.

b. Средняя ошибка аппроксимации $\bar{A}$ характеризует степень отклонения расчетных значений $\widehat{y_t}$от фактических y_t:

$$\bar{A} = \frac{1}{n}\sum_{t=1}^{n}\frac{|y_t - \widehat{y_t}|}{y_t}$$

Ошибка аппроксимации в пределах 5 – 10% свидетельствует о хорошем подборе модели к исходным данным.

Для прогнозов высокой точности $\bar{A}$ < 10%, хорошей – 10% < $\bar{A}$ < 20%, удовлетворительной – 20% < $\bar{A}$ < 50%, неудовлетворительной – $\bar{A}$ >50%.

c. Коэффициент детерминации показывает, какая доля изменчивости временного ряда объясняется построением модели:

$$R^2 = \frac{\sum_{t=1}^{n}(\hat{Y}_t - \bar{Y})^2}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}$$

Коэффициент детерминации принимает значения от 0 до 1. Чем ближе значение коэффициента к 1, тем лучше модель описывает поведение временного ряда, чем ближе к 0, делаем вывод простое среднее приближает лучше.

Пример 8. Имеется временной ряд y_t, $t = \overline{1, 47}$, содержащий данные об объеме контейнерных перевозок грузов железнодорожным транспортом, млн тонн (таб. 27).

Таблица 27

Исходные данные примера 8

t	y_t	t	y_t
1	10	25	39
2	13	26	37
3	19	27	43
4	24	28	50
5	22	29	40
6	22	30	48
7	26	31	48
8	23	32	49
9	19	33	46
10	17	34	47
11	20	35	38
12	22	36	47
13	18	37	48
14	23	38	54
15	29	39	56
16	27	40	56
17	29	41	56
18	25	42	57
19	31	43	59
20	33	44	54
21	35	45	55
22	46	46	57
23	32		
24	46		

Требуется:

1. Оценить параметры линейного тренда с помощью метода наименьших квадратов;
2. Верифицировать модель;
3. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}95$;
4. Оценить качество полученных результатов с помощью средней относительной ошибки аппроксимации.

1 шаг. Построить график изменения уровней временного ряда и добавить линию тренда (рис. 13).

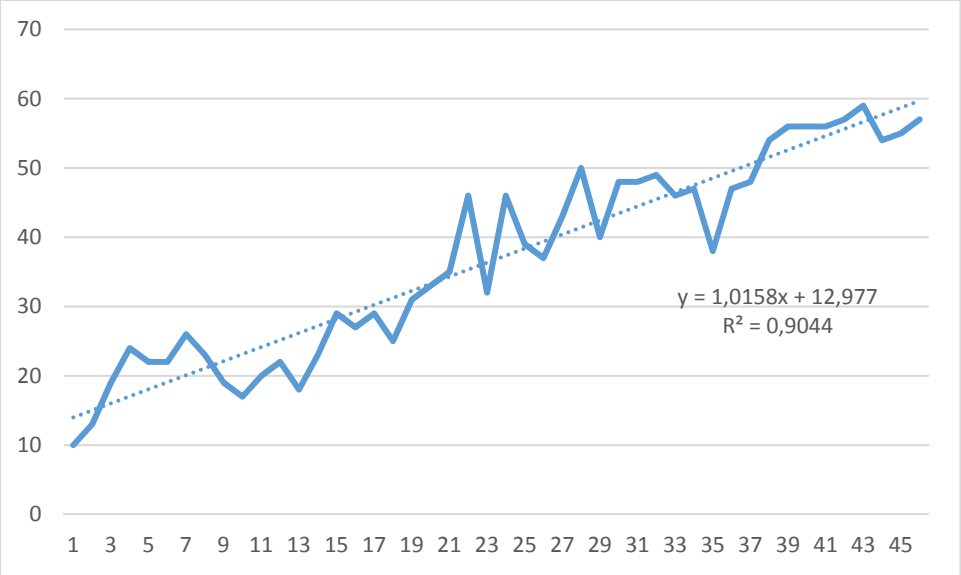


Рис. 13. Визуальное представление временного ряда пример 23

Коэффициент детерминации $R^2 = 0{,}9044$, что показывает 90,44% вариабельности уровней временного ряда определяется линейным влиянием *t*.

2 шаг. Построить линейную модель временного ряда и оценить ее параметры методом наименьших квадратов, используя команду **ДАННЫЕ/АНАЛИЗ ДАННЫХ/РЕГРЕССИЯ** (рис. 14).

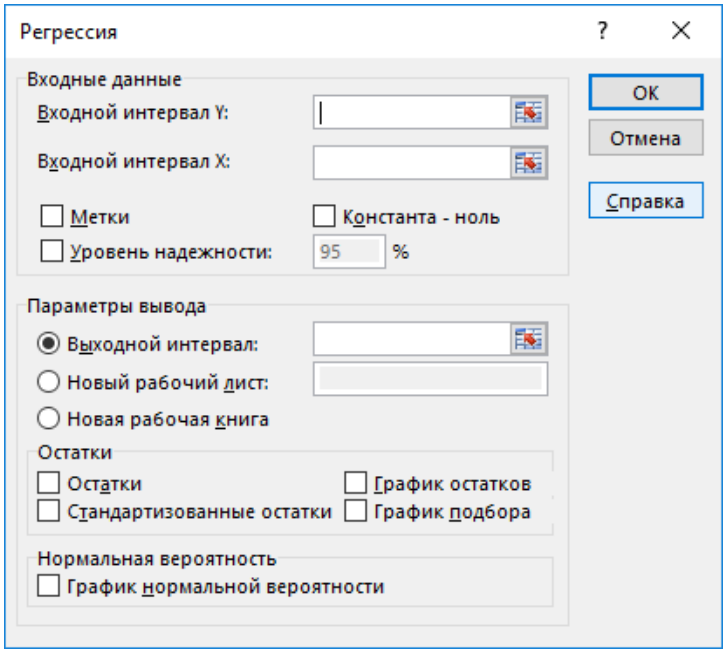


Рис. 14. Диалоговое окно команды ДАННЫЕ/АНАЛИЗ ДАННЫХ/РЕГРЕССИЯ

– Входной интервал Y — диапазон ячеек, содержащий уровни временного ряда:

– Входной интервал X — диапазон ячеек, содержащий значения t;

–Метки — установить флажок, если первые строки диапазонов данных содержат имена этих данных (заголовки);

– Константа-ноль — установить флажок, если строится регрессия без постоянной $y=bx$;

–Уровень надежности — установить флажок, если задается уровень надежности не 95% (γ=0,95), а другой, например, 99%;

– Параметры вывода — выбирать одно из мест расположения выводимых результатов:

- Выходной интервал – для помещения результатов на текущем рабочем листе, указать ячейку этого листа, начиная с которой будет выводиться результат;
- "Новый рабочий лист" – для расположения результатов на новом рабочем листе;
- "Новая книга" – для помещения результатов в новую рабочую книгу.

– Остатки — для выдачи прогнозов $\hat{y}_t$ и остатков регрессии $e_t = y_t - \hat{y}_t$;

– Стандартизованные остатки — для вывода нормированных остатков e_t/S;

График нормальной вероятности — для вывода таблицы, в которой указывается какими персентилями являются наблюдаемые значения зависимой переменной y, и построения соответствующего графика;

– График остатков — для вывода точечной диаграммы остатков e_t;

– График подбора — для вывода наложенных на диаграмму рассеяния точек $(x_t, \hat{y}_t)$ линии регрессии $\hat{y}_t = \hat{a} + \hat{b}x_t$.

Результат выполнения команды приведен на рисунке 15.

ВЫВОД ИТОГОВ

Регрессионная статистика	
Множественный R	0,950993
R-квадрат	0,904387
Нормированный R-квадрат	0,902214
Стандартная ошибка	4,483338
Наблюдения	46

Дисперсионный анализ

	df	*SS*	*MS*	*F*	*Значимость F*
Регрессия	1	8365,521	8365,521	416,1885	4,68E-24
Остаток	44	884,4139	20,10032		
Итого	45	9249,935			

	Коэффициенты	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*	*Нижние 95,0%*	*Верхние 95,0%*
Y-пересечение	12,97681	1,343918	9,655955	1,94E-12	10,26832	15,6853	10,26832	15,6853
t	1,015788	0,049792	20,4007	4,68E-24	0,915439	1,116137	0,915439	1,116137

ВЫВОД ОСТАТКА

Наблюдение	*Предсказанное yt*	*Остатки*
1	13,9926	-3,9926
2	15,00839	-2,00839
3	16,02418	2,975825
4	17,03996	6,960037
5	18,05575	3,944249
6	19,07154	2,928461
7	20,08733	5,912673
8	21,10311	1,896886
9	22,1189	-3,1189
10	23,13469	-6,13469
11	24,15048	-4,15048
12	25,16627	-3,16627
13	26,18205	-8,18205
14	27,19784	-4,19784
15	28,21363	0,786371
16	29,22942	-2,22942
	…	
37	50,56096	-2,56096
38	51,57675	2,42325
39	52,59254	3,407462
40	53,60833	2,391674
41	54,62411	1,375887
42	55,6399	1,360099
43	56,65569	2,344311
44	57,67148	-3,67148
45	58,68726	-3,68726
46	59,70305	-2,70305

Рис. 15. Результат использования команды Данные/Анализ данных/Регрессия

Таблица «Регрессионная статистика» содержит:

Множественный R – множественный коэффициент корреляции между y_t и $\hat{y}_t$, $R = \sqrt{R^2}$;

- R -квадрат – коэффициент детерминации $R^2 = 1 - \frac{SS_{ост}}{SS_{общ}}$;
- Нормированный R -квадрат – скорректированный коэффициент детерминации $\overline{R}^2 = 1 - (1 - R^2)\frac{k-1}{n-k}$, где k – число коэффициентов в модели;
- Стандартная ошибка – оценка среднеквадратического отклонения ошибок ε_i регрессии, т.е. $S = \sqrt{S^2} = \sqrt{\frac{1}{n-2}\sum_{t=1}^{n} e_t^2}$;

- Наблюдений – объем выборки *n*.

Таблица «Дисперсионный анализ», где:

- df – число степеней свободы;
- SS – сумма квадратов;
- F – вычисленное значение критерия Фишера (F-статистики);
- Значимость F – уровень значимости, при котором вычисленное значение критерия Фишера является критической точкой распределения Фишера. Нулевая гипотеза о незначимости (H_0: $b=0$) уравнения регрессии отклоняется, если это значение меньше заданного уровня значимости.

Третья таблица содержит МНК-оценки коэффициентов регрессии, их стандартные ошибки, значения t-статистик для проверки нулевых гипотез H_0: $a=0$ и H_0: $b=0$, P-вероятности и границы доверительных интервалов коэффициентов регрессии для заданных уровней надежности.

В строке с именем "Y-пересечение" приводятся:

- оценка $\hat{a}$ коэффициента *a*;
- ее стандартная ошибка $S_{\hat{a}} = \sqrt{\hat{D}(\hat{a})}$;
- вычисленное значение t-статистики $t = \hat{a}/S_{\hat{a}}$;
- P-значение, – вероятность того, что случайная величина распределенная по закону $t\ (n\text{-}2)$ примет значение по абсолютной величине больше, чем модуль вычисленного значения t-статистики, т.е. P-значение это уровень значимости, при котором вычисленное значение t-статистики является критической точкой, следовательно, нулевая гипотеза H_0: $a=0$ отклоняется, если P-значение меньше заданного уровня значимости, и принимается в противном случае;
- нижняя и верхняя границы 95%-ого доверительный интервал для *a*;
- нижняя и верхняя границы 99%-ого доверительного интервала для *a*.

В строке с именем "X" приводятся аналогичные данные для коэффициента *b* уравнения регрессии.

Таблица "Вывод остатка" содержит порядковые номера наблюдений ($t=1,2,\ldots,n$), предсказанные (прогнозные) значения среднего зависимой переменной $\hat{y}_t = \hat{a} + \hat{b}x_t$. и остатки регрессии $e_t = y_t - \hat{y}_t$..

Графики подбора и остатков переместить на рабочем листе на подходящие места. Используя средства Excel, изменить, при необходимости, их масштабы, надписи, легенду, форматы осей, шрифты. На графике подбора выводится диаграмма рассеяния и точки $(x_t, \hat{y}_t)$ линии регрессии $\hat{y}_t = \hat{a} + \hat{b}x_t$.

На графике остатков представлены остатки e_t для наблюдаемых значений x_t.

Таким образом, выполнение команды **ДАННЫЕ/АНАЛИЗ ДАННЫХ/РЕГРЕССИЯ** позволяет получить:

- уравнение регрессии, $\hat{y} = 12{,}97681 + 1{,}015788 \cdot x$;
- оценку среднеквадратического отклонения ошибок регрессии S=4,483338 и оценку дисперсии ошибок S^2 =20,1003
- 95%-е доверительные интервалы для коэффициентов регрессии 10,26832<*a*<15,6853 и 0,915439<*b*<1,116137;

• вычисленное значение t-статистики для коэффициента *a*, t = 9,655955, и ее P-значение, равное $1{,}94 \cdot 10^{12}$. Так как P-значение значительно меньше заданных уровней значимости 0,05 и 0,01 отклоняем гипотезу H_0: $a=0$ о незначимости коэффициента *a* и принимаем альтернативную H_0: $a \neq 0$.

• вычисленное значение t-статистики для коэффициента *b*, t = 20,4007, и ее P-значение равное $4{,}68 \cdot 10^{24}$, что значительно меньше заданных уровней значимости 0,05 и 0,01, поэтому отклоняем гипотезу H_0: $b=0$ о незначимости коэффициента *b* и принимаем альтернативную H_0: $b \neq 0$;

• коэффициент детерминации R^2 = 0,904387, вычисленное значение F-статистики, F=426,1885, и ее уровень значимости, равный $4{,}68 \cdot 10^{24}$, что значительно меньше заданного уровня значимости 0,05, это позволяет отклонить нулевую гипотезу о незначимости коэффициента R^2 (о незначимости уравнения регрессии);

• линию регрессии, наложенную на диаграмму рассеяния;

• выборочный коэффициент корреляции, совпадающий со значением "Множественный R" таблицы "Регрессионная статистика", т.е. $\hat{r} = 0{,}950993$.

3 шаг. Рассчитать точечные и интервальные прогнозы, для этого заполнить таблицу вида

Таблица 28

Расчет прогнозов

Период	***Точечный прогноз***	***Нижняя граница***	***Верхняя граница***
n + 1			
n + 2			
n + 3			
n + 4			

Точечный прогноз получается подстановкой в уравнение регрессии

$$y=12{,}97681+1{,}015788*t$$

значения t, равного 47, 48, 49, 40 соответственно. Для определения нижних и верхних границ рассчитаем доверительный полуинтервал с доверительной вероятностью p = 0,95. $\Delta = t_{табл} \cdot \sigma$, где $t_{табл}$ определяется с помощью функции СТЬЮДЕНТ.ОБР.2Х(), а σ – СТАНДОТКЛОН() (таб. 29).

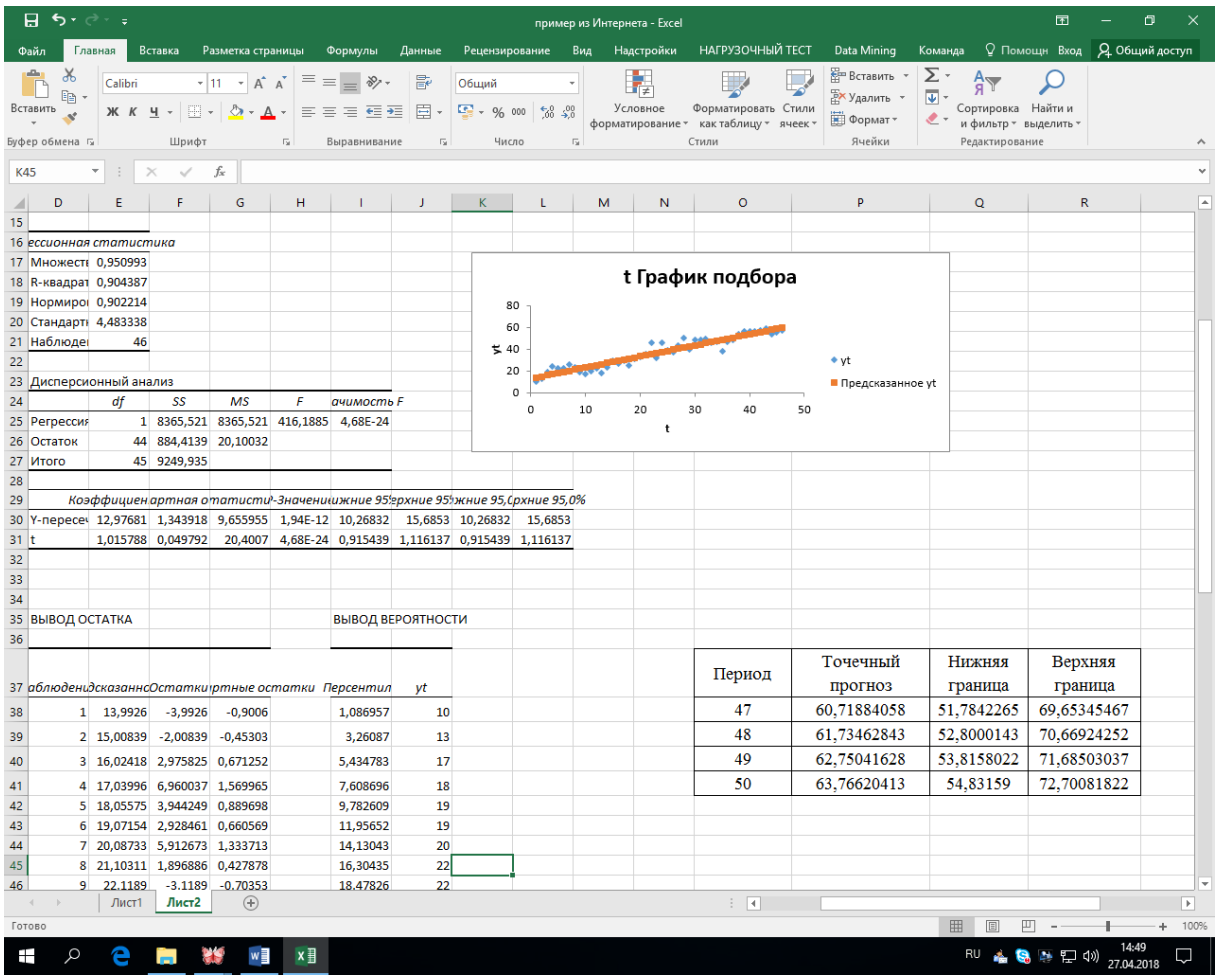
Таблица 29

Результаты расчета доверительного полуинтервала Δ

Наблюдение	*Предсказанное yt*	*Остатки*
1	13,9926	-3,9926
2	15,00839	-2,00839
3	16,02418	2,975825
4	17,03996	6,960037
5	18,05575	3,944249
6	19,07154	2,928461

Наблюдение	*Предсказанное yt*	*Остатки*
7	20,08733	5,912673
8	21,10311	1,896886
9	22,1189	-3,1189
10	23,13469	-6,13469
11	24,15048	-4,15048
12	25,16627	-3,16627
13	26,18205	-8,18205
14	27,19784	-4,19784
15	28,21363	0,786371
16	29,22942	-2,22942
	…	
37	50,56096	-2,56096
38	51,57675	2,42325
39	52,59254	3,407462
40	53,60833	2,391674
41	54,62411	1,375887
42	55,6399	1,360099
43	56,65569	2,344311
44	57,67148	-3,67148
45	58,68726	-3,68726
46	59,70305	-2,70305
σ=		4,433243
$t_{табл}$		2,015368
Δ		8,934614

На рабочем листе ввести таблицу в соответствии с образцом (таб. 28) и заполнить ячейки с точечными и интервальными прогнозами (рис. 16).



Период	Точечный прогноз	Нижняя граница	Верхняя граница
47	60,71884058	51,7842265	69,65345467
48	61,73462843	52,8000143	70,66924252
49	62,75041628	53,8158022	71,68503037
50	63,76620413	54,83159	72,70081822

Рис. 16. Расчет точечных и интервальных прогнозов

4 шаг. Рассчитать среднюю относительную ошибку аппроксимации по формуле:

$$\overline{A} = \frac{1}{n}\sum_{t=1}^{n}\frac{|y_t - \hat{y}_t|}{y_t}100\% ,$$

где $\hat{y}_t$ - расчетные (модельные) значения уровней ряда.

Ошибка аппроксимации ранжируется по значениям:

- прогнозов высокой точности $\overline{A} < 10\%$,
- хорошей – $10\% < \overline{A} < 20\%$,
- удовлетворительной – $20\% < \overline{A} < 50\%$,
- неудовлетворительной – $\overline{A} > 50\%$.

Целесообразно пропускать значения ряда, для которых $y_t = 0$.

Исходные данные для расчетов лучше брать из таблицы Вывод остатков результатов выполнения команды **ДАННЫЕ/АНАЛИЗ ДАННЫХ/РЕГРЕССИЯ**, где $\hat{y}_t$ - расчетные значения уровней ряда – предсказанные y_t, разность $y_t - \hat{y}_t$ – остатки (рис. 15). В результате расчетов (таб. 30) получить значение 11% (хорошее качество прогноза).

Таблица 30

Результат расчета средней ошибки аппроксимации

t	y_t	y_t-расчетные	остатки
1	10	13,9926	0,39926
2	13	15,00839	0,154491
3	19	16,02418	0,156622
4	24	17,03996	0,290002
5	22	18,05575	0,179284
6	22	19,07154	0,133112
7	26	20,08733	0,227411
8	23	21,10311	0,082473
9	19	22,1189	0,164153
…	…	…	…
32	49	45,48202	0,071795
33	46	46,49781	0,010822
34	47	47,5136	0,010928
35	38	48,52939	0,277089
36	47	49,54517	0,054153
37	48	50,56096	0,053353
38	54	51,57675	0,044875
39	56	52,59254	0,060848
40	56	53,60833	0,042708
41	56	54,62411	0,024569
42	57	55,6399	0,023861
43	59	56,65569	0,039734
44	54	57,67148	0,06799
45	55	58,68726	0,067041
46	57	59,70305	0,047422
$t_{табл}$		2,015368	11%
Ошибка аппроксимации		2,015368	

Список рекомендуемой литературы

1. Афанасьев В. Н., Юзбашев М. М. Анализ временных рядов и прогнозирование. допущено М-вом образования РФ : учеб. для вузов / В. Н. Афанасьев, М. М. Юзбашев. – М. : Финансы и статистика, 2001. – 227 с.

2. Головченко В. Б., Васильев С. Н. Прогнозирование временных рядов по разнородной информации / В. Б. Головченко. – Новосибирск : Наука, 1999. – 88 с.

3. Музыко Е.И. Экономическое прогнозирование : учебно-метод. пособие / Е.И. Музыко. – Министерство образования и науки Российской Федерации, Новосибирский государственный технический университет. – Новосибирск: НГТУ, 2015. – 240 с. : схем., табл. – ISBN 978-5-7782-2701-9 ; То же [Электронный ресурс]. – Режим доступа: http://biblioclub.ru/index.php?page=book&id=438414 (31.05.2016).

4. Банкротство предприятия: анализ, учет и прогнозирование : учеб. Пособие / К. В. Балдин и др. – М. : Дашков и К, 2007. – 375 с.

5. Бизнес-прогнозирование. Business Forecasting. Business Forecasting. 7-е изд. / Дж. Э. Ханк, А. Дж. Райтс, Д. У. Уичерн. – М. : Вильямс, 2003. – 653 с.

6. Писарева О. М. Методы прогнозирования развития социально-экономических систем. допущено Советом УМО вузов РФ : учеб. Пособие / О. М. Писарева. – М. : Высш. шк., 2007. – 591 с.

7. Боровиков В.Н., Ивченко Г.И. Прогнозирование в системе STATISTICA в среде Windows. Основы теориии и интенсивная практика на компьютере. рек. М-вом образования РФ : учеб. пособие / В. П. Боровиков, Г.И. Ивченко. – М. : Финансы и статистика, 2000. – 382 с.

8. Прогнозирование и планирование экономики : учеб. пособие для вузов. допущено М-вом образования Респ. Беларусь / В. И. Борисевич и др. – Минск : Экоперспектива, 2000. – 431 с.

9. Головченко В. Б. Прогнозирование с использованием разнородной информации : учеб. пособие / В.Б. Головченко. – Иркутск : Изд-во БГУЭП, 2005. – 71 с.

10. Зехин В.А. Практикум по многомерным статистическим методам: учебное пособие / В.А. Зехин, В.С. Мхитарян, С.А. Айвазян. – 1-е изд. – М. : Московский государственный университет экономики, статистики и информатики, 2003. – 76 с.; То же [Электронный ресурс]. – Режим доступа: http://biblioclub.ru/index.php?page =book&id=90409 (31.05.2016).

ПРИЛОЖЕНИЕ 1

Сглаживание временного ряда методом простой скользящей средней средствами Excel

В Excel реализован алгоритм простой скользящей средней при нечетном интервале сглаживания. Для сглаживания уровней временного ряда методом скользящей средней рекомендуется выполнить следующие шаги.

1 шаг. На рабочем листе сформировать исходный ряд (рекомендуется разместить значения ряда в столбец);

2 шаг. Выполнить команду **ДАННЫЕ/АНАЛИЗ ДАННЫХ/СКОЛЬЗЯЩЕЕ СРЕДНЕЕ;**

3 шаг. В открывшемся диалоговом окне «Скользящее среднее» (рис. 17) задать параметры:

– в поле «*Входной интервал*» – диапазон ячеек, содержащий исходный ряд;

– в поле «*Интервал*» – значение периода сглаживания;

– в поле «*Выходной интервал*» – адрес ячейки, начиная с которой будет выводиться результат.

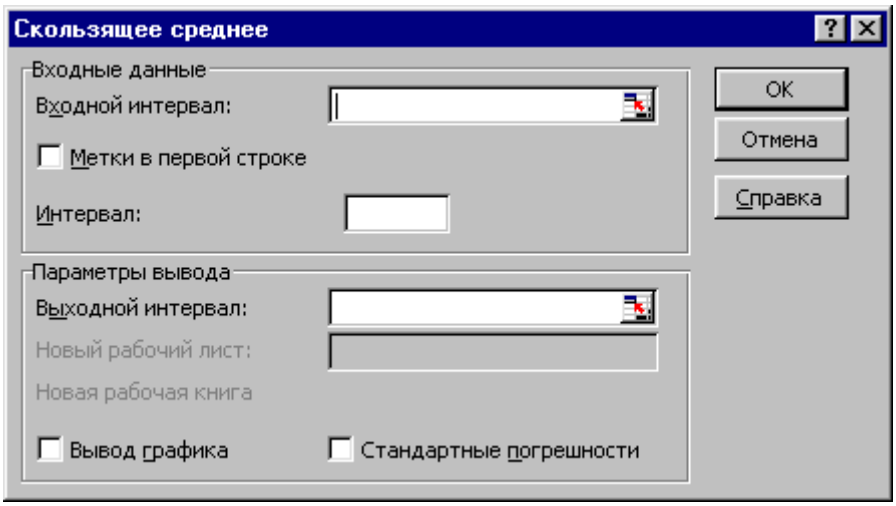


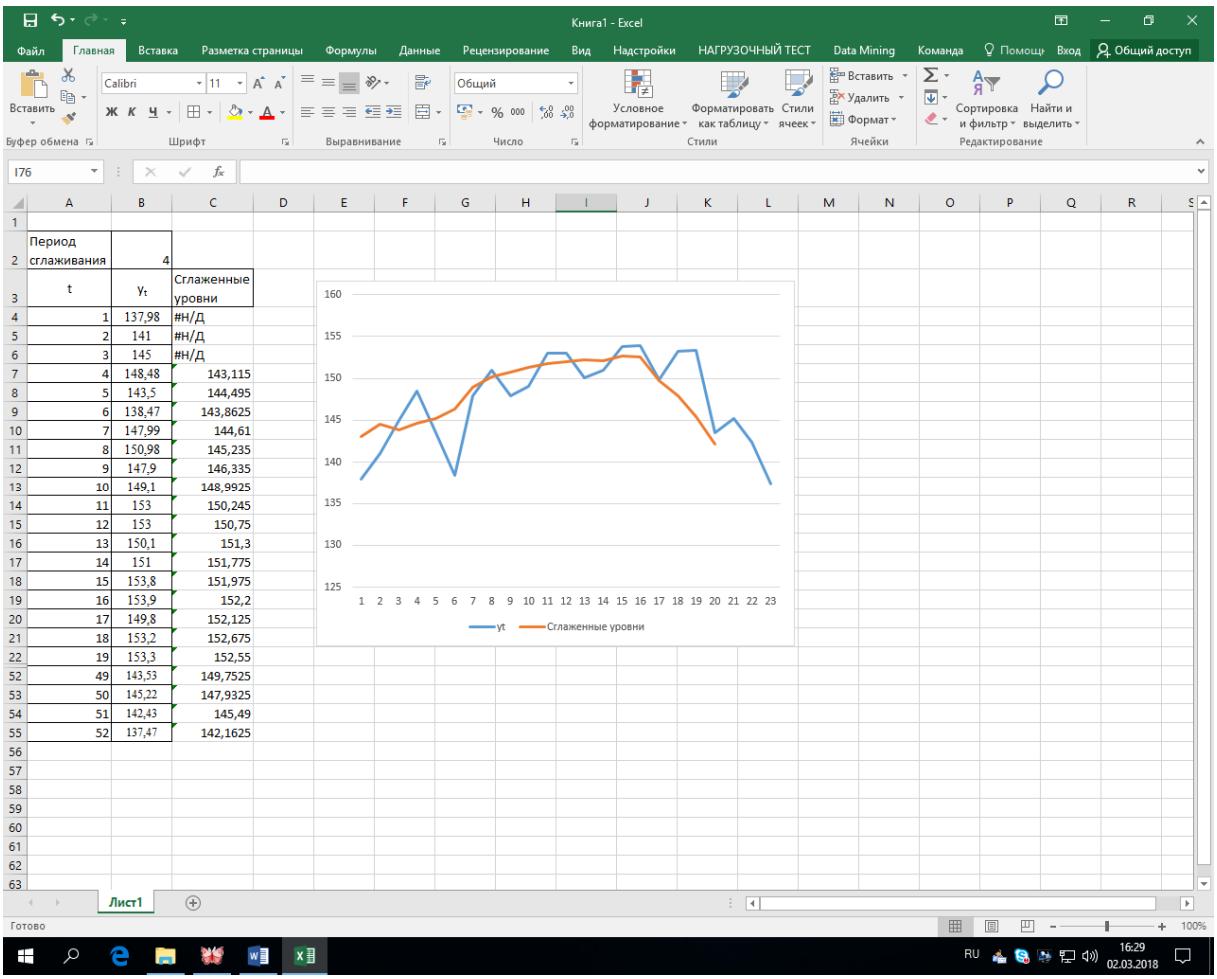
Рис. 17. Диалоговое окне «Скользящее среднее»

Результатом выполнения команды является столбец сглаженных уровней ряда. Причем, для первых и последних m-1 (в зависимости от периода сглаживания) уровней скользящее среднее не подсчитывается.

Пример 9. Имеется временной ряд y_t, $t = \overline{1, 52}$, содержащий данные о добыче нефти в СНГ, млн. т. (рис. 18). Требуется:

1. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания $m = 4$) и привести график сравнения исходных и сглаженных уровней временного ряда.

Результат выполнения задания приведен на рис. 18.



Период сглаживания	4	
t	y_t	Сглаженные уровни
1	137,98	#Н/Д
2	141	#Н/Д
3	145	#Н/Д
4	148,48	143,115
5	143,5	144,495
6	138,47	143,8625
7	147,99	144,61
8	150,98	145,235
9	147,9	146,335
10	149,1	148,9925
11	153	150,245
12	153	150,75
13	150,1	151,3
14	151	151,775
15	153,8	151,975
16	153,9	152,2
17	149,8	152,125
18	153,2	152,675
19	153,3	152,55
49	143,53	149,7525
50	145,22	147,9325
51	142,43	145,49
52	137,47	142,1625

Рис. 18. Сглаживание временного ряда методом простой скользящей средней

ПРИЛОЖЕНИЕ 2

Сглаживание временного ряда методом экспоненциального сглаживания средствами Excel

В Excel реализован алгоритм экспоненциального сглаживания с заданным периодом сглаживания. Для сглаживания уровней временного ряда методом экспоненциального сглаживания рекомендуется выполнить следующие шаги.

1 шаг. На рабочем листе сформировать исходный ряд (рекомендуется разместить значения ряда в столбец);

2 шаг. Выполнить команду **ДАННЫЕ/АНАЛИЗ ДАННЫХ/ЭКСПОНЕНЦИАЛЬНОЕ СГЛАЖИВАНИЕ;**

3 шаг. В открывшемся диалоговом окне «Экспоненциальное сглаживание» (рис. 19) задать параметры:

– в поле «Входной интервал» – диапазон, содержащий исходный ряд;
– в поле «Фактор затухания» – параметр сглаживания;
– в поле «Выходной интервал» – адрес ячейки, начиная с которой будет выводиться результат.

Результатом выполнения команды является столбец сглаженных значений уровней ряда. При этом первое значение уровня сглаженного ряда не подсчитывается.

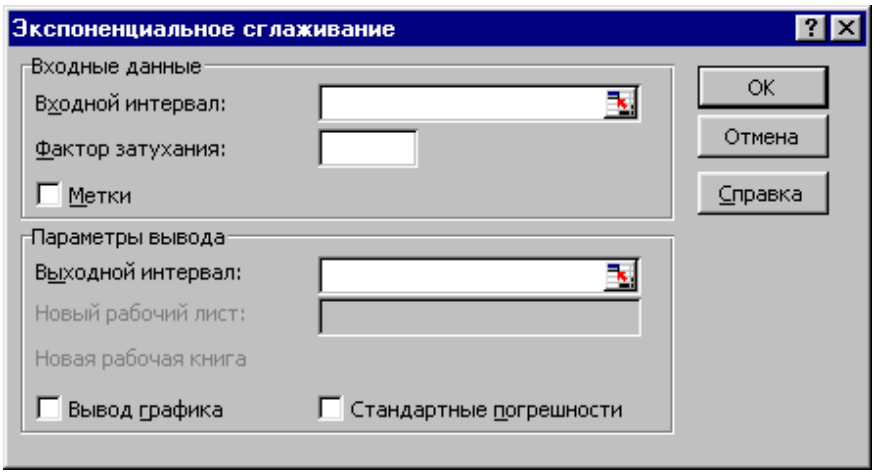
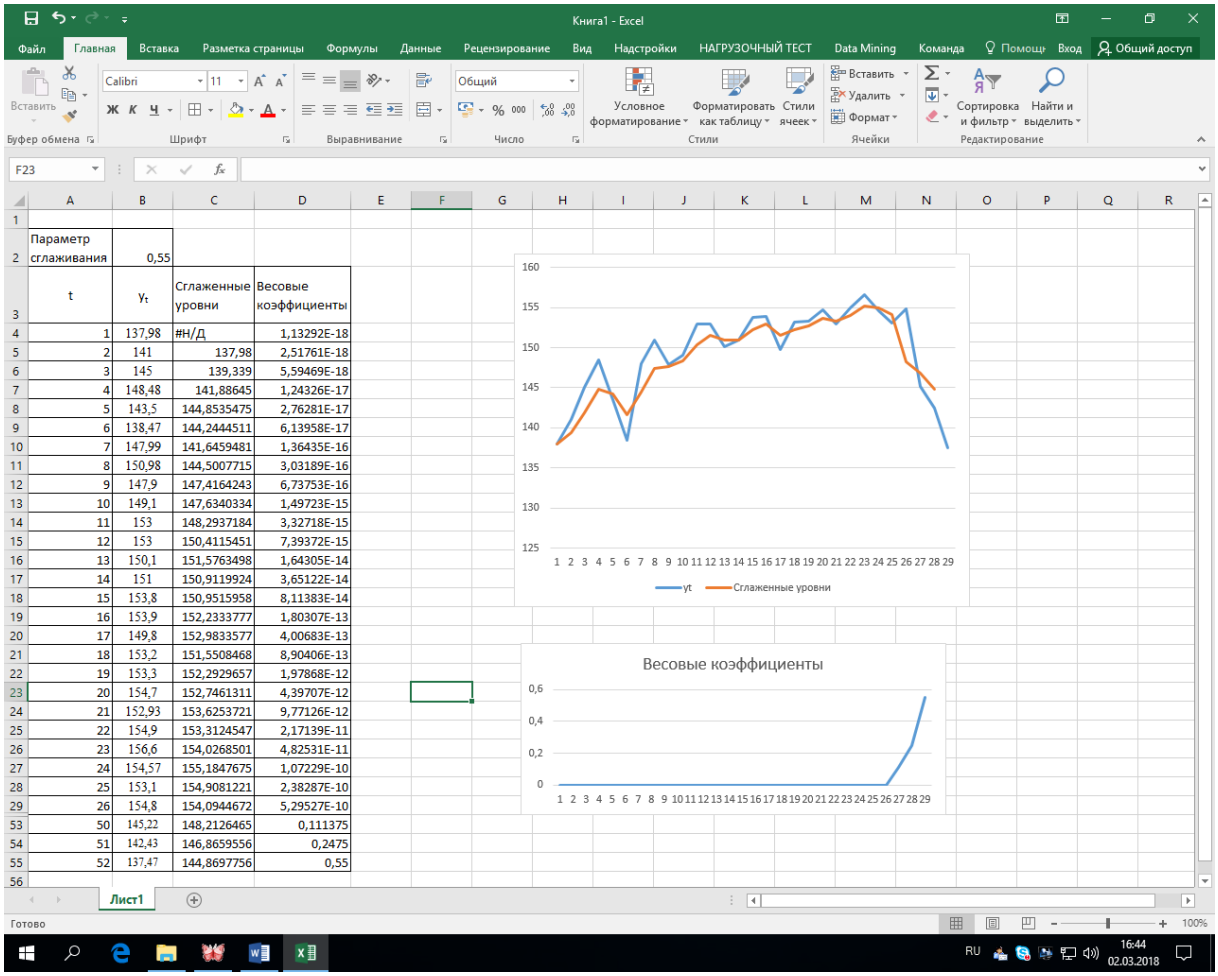


Рис. 19. Диалоговое окно команды
ДАННЫЕ/АНАЛИЗ ДАННЫХ/ЭКСПОНЕНЦИАЛЬНОЕ СГЛАЖИВАНИЕ

Пример 10. Имеется временной ряд y_t, $t = \overline{1, 52}$, содержащий данные о добыче нефти в СНГ, млн. т. (рис. 20). Требуется:

1. Сгладить временной ряд с помощью экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}55$).
2. Рассчитать весовые коэффициенты и привести график сравнения исходных и сглаженных уровней временного ряда и график весовых коэффициентов.
3. Результат выполнения задания приведен на рис. 20.



Параметр сглаживания	0,55		
t	y_t	Сглаженные уровни	Весовые коэффициенты
1	137,98	#Н/Д	1,13292E-18
2	141	137,98	2,51761E-18
3	145	139,339	5,59469E-18
4	148,48	141,88645	1,24326E-17
5	143,5	144,8535475	2,76281E-17
6	138,47	144,2444511	6,13958E-17
7	147,99	141,6459481	1,36435E-16
8	150,98	144,5007715	3,03189E-16
9	147,9	147,4164243	6,73753E-16
10	149,1	147,6340334	1,49723E-15
11	153	148,2937184	3,32718E-15
12	153	150,4115451	7,39372E-15
13	150,1	151,5763498	1,64305E-14
14	151	150,9119924	3,65122E-14
15	153,8	150,9515958	8,11383E-14
16	153,9	152,2333777	1,80307E-13
17	149,8	152,9833577	4,00683E-13
18	153,2	151,5508468	8,90406E-13
19	153,3	152,2929657	1,97868E-12
20	154,7	152,7461311	4,39707E-12
21	152,93	153,6253721	9,77126E-12
22	154,9	153,3124547	2,17139E-11
23	156,6	154,0268501	4,82531E-11
24	154,57	155,1847675	1,07229E-10
25	153,1	154,9081221	2,38287E-10
26	154,8	154,0944672	5,29527E-10
50	145,22	148,2126465	0,111375
51	142,43	146,8659556	0,2475
52	137,47	144,8697756	0,55

Рис. 20. Сглаживание временного ряда методом экспоненциального сглаживание

ПРИЛОЖЕНИЕ 3

Преобразование нелинейной функции к линейному виду

Метод наименьших квадратов используется для получения оценок параметров только линейных функций. Функции, не линейные по параметрам и/или переменным, необходимо предварительно линеаризовать. Простейшие методы линеаризации – введение новой переменной или логарифмирование. В качестве примера рассмотрим степенную функцию вида

$$y_t = a_0 t^{a_1} \varepsilon_t . \qquad (*)$$

1. Имеется временной ряд $y_t, \quad t = \overline{1, n}$. Принято решение, что данный ряд целесообразно аппроксимировать степенной функцией.

2. Прологарифмировав исходный ряд, имеем новый временной ряд вида $y'_t = \ln y_t, \quad t = \overline{1, n}$, а модель (*) принимает вид

$$y'_t = \ln a_0 + a_1 \ln t + \ln \varepsilon_t . \qquad (**)$$

3. Оцениваются параметры уравнения (**). Следует обратить внимание, что в качестве аргумента функции используется не *t*, а ln *t*. В результате применения процедуры имеем ln a_0 и a_1.

4. Для перехода от ln a_0 к a_0 вычислить $a_0 = e^{\ln a0}$.

Пример 11. Имеется временной ряд $y_t, \quad t = \overline{1, 35}$, содержащий данные о расходf[на жилье $\mathbf{y_t}$ в млрд. долл. США в ценах 2017 г. (таб. 31). Требуется:

1. Построить нелинейную модель тренда, например, $y_t = a_0 t^{a_1} \varepsilon_t$ где ε_t – случайный мультипликативный член, изменяющий $y_t = a_0 t^{a_1}$ в случайной пропорции.
2. Оценить точность и надежность оценок параметров модели.
3. Провести верификацию модели.
4. Построить прогнозы среднего значения временного ряда.

Этапы выполнения задания.

1. Провести преобразование нелинейной модели к линейному, относительны параметрам, виду. Используя заданную выборку значений, вычислить значения новых переменных $\ln y_t$ и $\ln t$. Построить для них диаграмму рассеяния и визуально оценить возможность построения линейной модели в новых переменных.

2. Найти МНК – оценки коэффициентов парной линейной регрессии для новых переменных. Построить найденную линию регрессии, наложенную на диаграмму рассеяния.

3. Найти оценки среднеквадратических отклонений (стандартные ошибки) оценок коэффициентов линейной регрессии и оценку дисперсии ошибки линейной регрессии. Построить 95%-ые доверительные интервалы для коэффициентов линейной регрессии.

4. Провести верификацию оцененной линейной модели: проверить гипотезы о значимости коэффициентов регрессии при 5%-ом уровне значимости; вычислить коэффициент детерминации R^2 и проверить гипотезу о его значимости при 5%-ом уровне значимости.

5. Используя полученные результаты по преобразованной (линейной) модели, найти оценки параметров исходной нелинейной модели и записать эту модель. Указать доверительные интервалы для ее параметров.

6. Построить точечные прогнозы среднего зависимой переменной по построенной нелинейной модели для выборочных значений t. Построить линию регрессии. Вычислить индекс корреляции (корреляционное отношение) и среднюю ошибку аппроксимации.

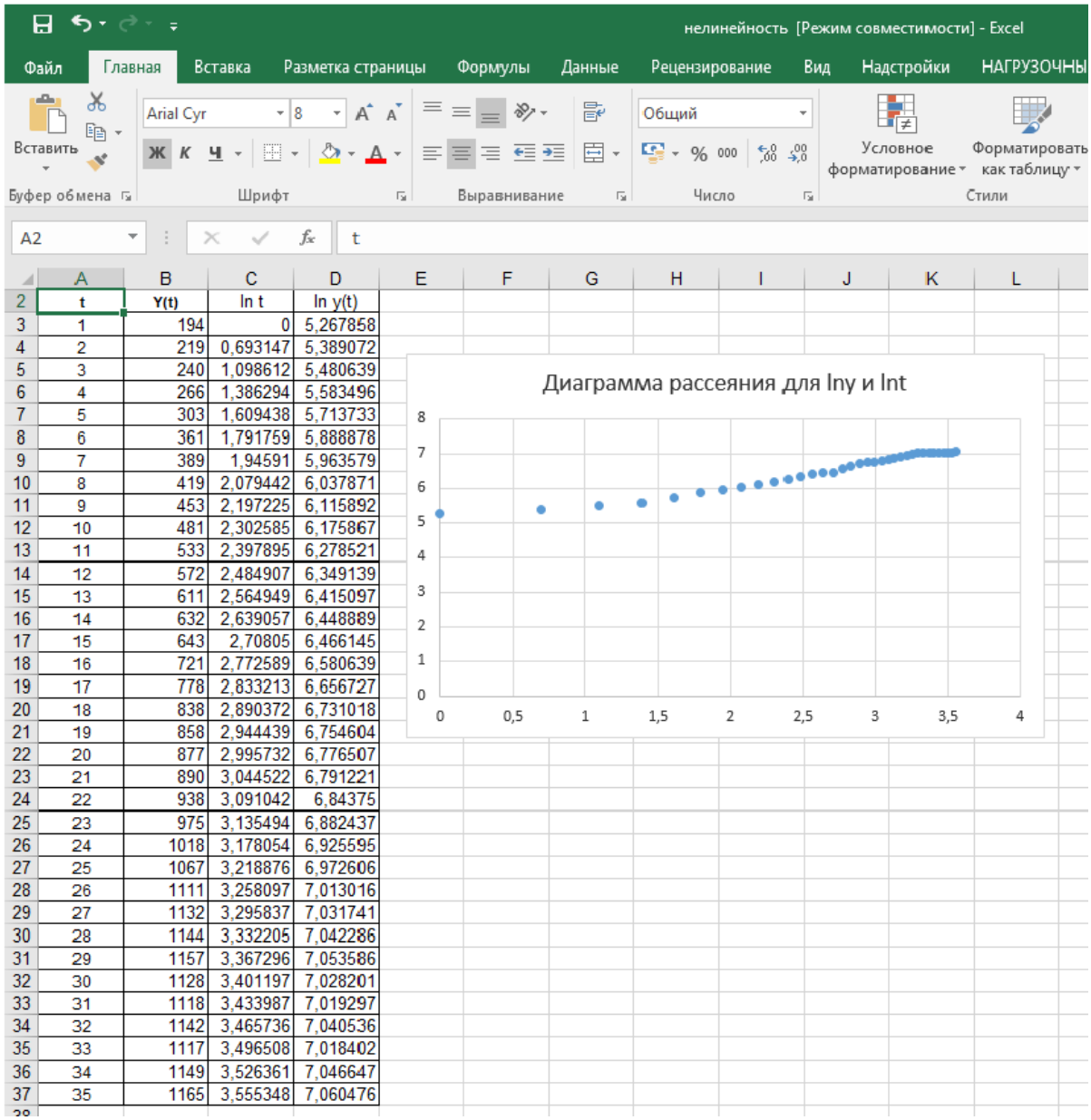
7. Дать интерпретацию построенной нелинейной модели и общее заключение о ней.

Таблица 31

Исходные данные примера 11

t	**y_t**	**t**	**y_t**
1	194	18	838
2	219	19	858
3	240	20	877
4	266	21	890
5	303	22	938
6	361	23	975
7	389	24	1018
8	419	25	1067
9	453	26	1111
10	481	27	1132
11	533	28	1144
12	572	29	1157
13	611	30	1128
14	632	31	1118
15	643	32	1142
16	721	33	1117
17	778	34	1149
		35	1165

1этап. Преобразование модели и определение новых переменных/ Прологарифмировать по натуральному основанию исходную нелинейную модель $y_t = a_0 t^{a_1} \varepsilon_t$, получить $lny_t = lna_0 + a_1 lnt + ln\varepsilon_t$. Определить новые переменные ln y_t и ln t. Вычислить значения новых переменных и построить диаграмму рассеяния в новых переменных (рис. 21).

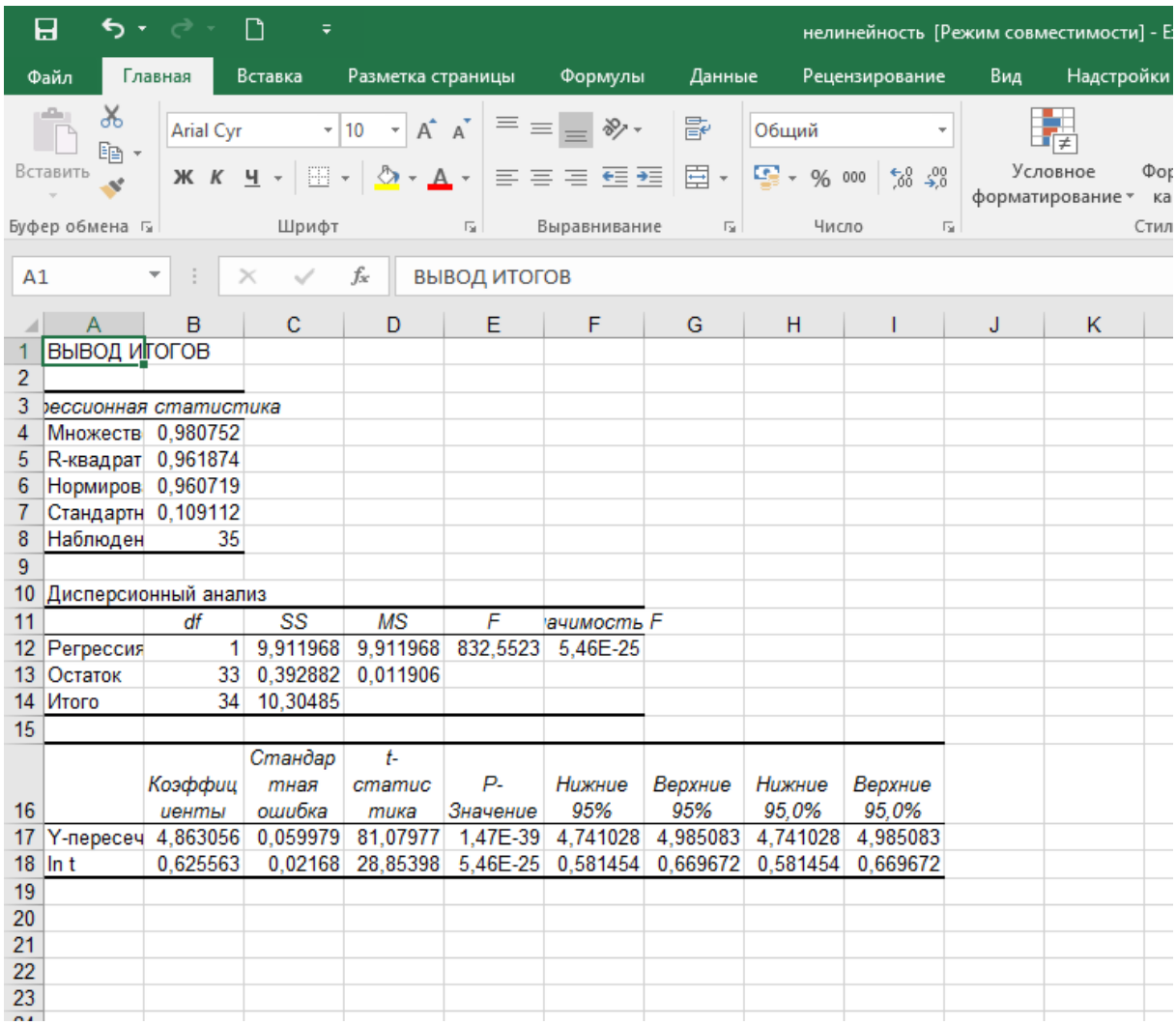


	A	B	C	D
2	t	Y(t)	ln t	ln y(t)
3	1	194	0	5,267858
4	2	219	0,693147	5,389072
5	3	240	1,098612	5,480639
6	4	266	1,386294	5,583496
7	5	303	1,609438	5,713733
8	6	361	1,791759	5,888878
9	7	389	1,94591	5,963579
10	8	419	2,079442	6,037871
11	9	453	2,197225	6,115892
12	10	481	2,302585	6,175867
13	11	533	2,397895	6,278521
14	12	572	2,484907	6,349139
15	13	611	2,564949	6,415097
16	14	632	2,639057	6,448889
17	15	643	2,70805	6,466145
18	16	721	2,772589	6,580639
19	17	778	2,833213	6,656727
20	18	838	2,890372	6,731018
21	19	858	2,944439	6,754604
22	20	877	2,995732	6,776507
23	21	890	3,044522	6,791221
24	22	938	3,091042	6,84375
25	23	975	3,135494	6,882437
26	24	1018	3,178054	6,925595
27	25	1067	3,218876	6,972606
28	26	1111	3,258097	7,013016
29	27	1132	3,295837	7,031741
30	28	1144	3,332205	7,042286
31	29	1157	3,367296	7,053586
32	30	1128	3,401197	7,028201
33	31	1118	3,433987	7,019297
34	32	1142	3,465736	7,040536
35	33	1117	3,496508	7,018402
36	34	1149	3,526361	7,046647
37	35	1165	3,555348	7,060476

Рис. 21. Расчет значений новых переменных и построение диаграммы

Точки диаграммы рассеяния близки некоторой прямой, следовательно, можно рассматривать линейную модель.

2 – 4 этапы. Оценка преобразованной модели получается при использовании команды **ДАННЫЕ/АНАЛИЗ ДАННЫХ/РЕГРЕССИЯ** и получении оценок неизвестных параметров (рис.22). Затем построить диаграмму рассеяния для новых переменных, добавить линейный тренд и вывести уравнение тренда и коэффициент детерминации R^2 (рис. 23).



	A	B	C	D	E	F	G	H	I
1	ВЫВОД ИТОГОВ								
2									
3	ессионная статистика								
4	Множеств	0,980752							
5	R-квадрат	0,961874							
6	Нормиров	0,960719							
7	Стандартн	0,109112							
8	Наблюден	35							
9									
10	Дисперсионный анализ								
11		df	SS	MS	F	ачимость F			
12	Регрессия	1	9,911968	9,911968	832,5523	5,46E-25			
13	Остаток	33	0,392882	0,011906					
14	Итого	34	10,30485						
15									
16		Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
17	Y-пересеч	4,863056	0,059979	81,07977	1,47E-39	4,741028	4,985083	4,741028	4,985083
18	ln t	0,625563	0,02168	28,85398	5,46E-25	0,581454	0,669672	0,581454	0,669672

Рис. 22. Результаты оценивания модели с логарифмами от переменных

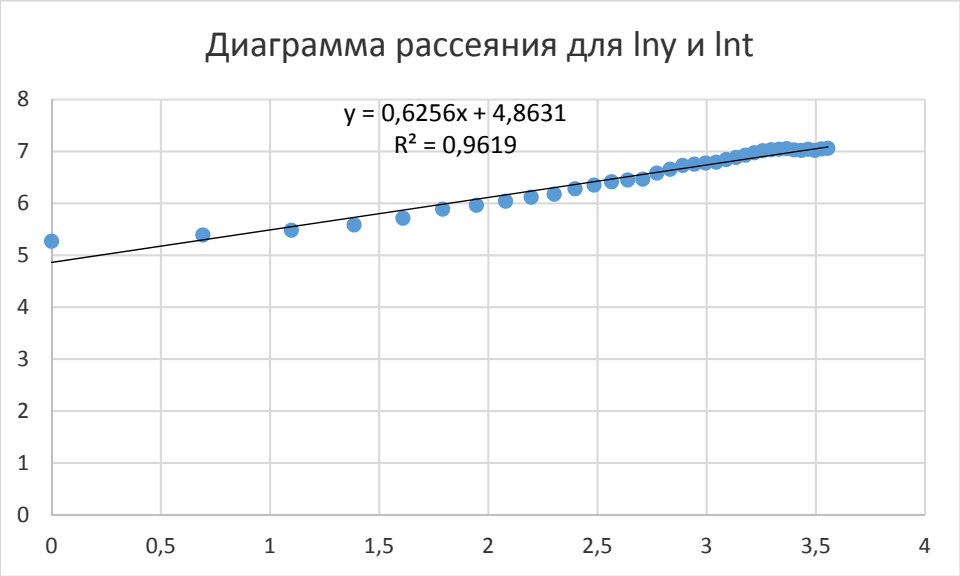


Рис. 23. Построение линии регрессии, наложенной на диаграмму рассеяния

Стандартные ошибки оценок параметров $S_{lna_0} = 0{,}059979$, $S_{a_1} = 0{,}02168$, а стандартная ошибка регрессии S=0,109112. 95%-е доверительные интервалы для lna_0 и a_1: $4{,}74108 < lna_0 < 4{,}985083$; $0{,}581454 < a_1 < 0{,}669672$.

Нулевые значения параметров не принадлежат этим доверительным интервалам, поэтому коэффициенты регрессии значимо отличаются от нуля при уровне значимости $\alpha = 0{,}05$, что также следует из малости P-значений t-статистик для параметров модели. Высокие значения коэффициента детерминации, $R^2 = 0{,}961874$, и F-статистики, $F = 832{,}553$, говорят о значимости оцененной модели (вычисленная "Значимость F"=5,46*10^{-25} существенно меньше $\alpha = 0{,}05$). При этом не нужно забывать, что полученные свойства оценок относятся к линейной модели в новых переменных, а не к исходной нелинейной модели. Привести сравнений исходных и расчетных по модели уровней временного ряда (рис. 24).

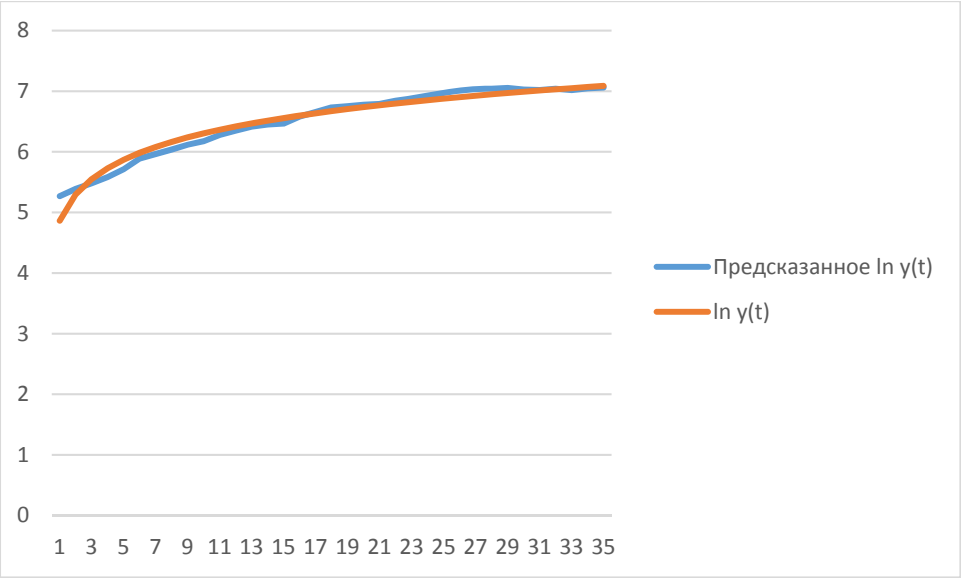


Рис. 24. Диаграмма сравнения исходных
и модельных уровней временного ряда для логарифмов

5 – 6 этапы. Переход к исходной нелинейной модели

Учитывая, что $\hat{a}_0 = e^{lna_0}$, а параметр a_1 в нелинейной модели совпадает с параметром a_1 в линейной модели для логарифмов переменных, получаем искомую нелинейную модель $y_t = e^{4{,}8631} t^{0{,}6256} = 129{,}4191 t^{0{,}6256}$.

95%-й доверительный интервал для параметра a_1 остается тем же самым $0{,}581454 < a_1 < 0{,}669672$, а для параметра a_0 имеет вид $114{,}552 = e^{4{,}741028} < a_0 < e^{4{,}985083} = 146{,}2158$.

Найдем расчетные значения временного ряда $\hat{y}_t$. Для этого воспользуемся уже вычисленными прогнозами lny_t, т.е. $\hat{y}_t = e^{lny_t}$. Также подсчитаем остатки нелинейной регрессии равные $y_t - \hat{y}_t$, рассчитаем ошибку аппроксимации для линейной модели и индекс корреляции для построенной нелинейной модели (рис. 25)

$$\eta = \sqrt{\frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}} = 0{,}91553109 .$$

Построим график, на котором сравниваются уровни исходного и рассчитанного по модели временного ряда (рис. 26).

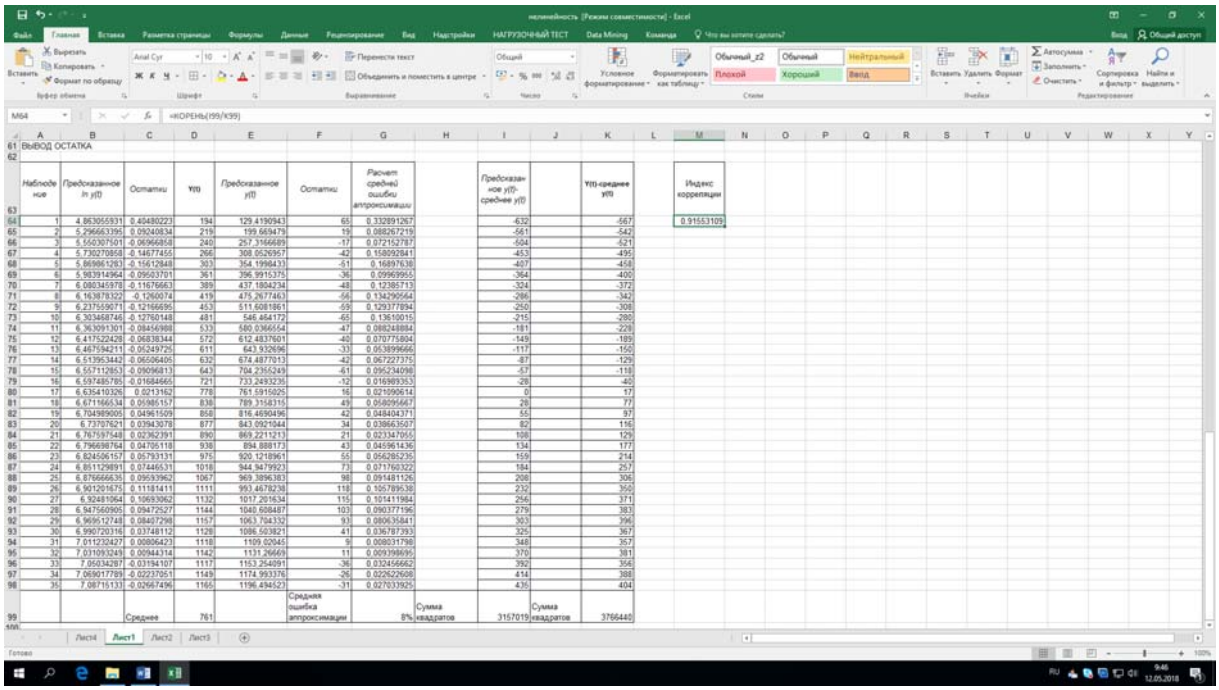


Рис. 25. Расчет ошибки аппроксимации и индекса корреляции

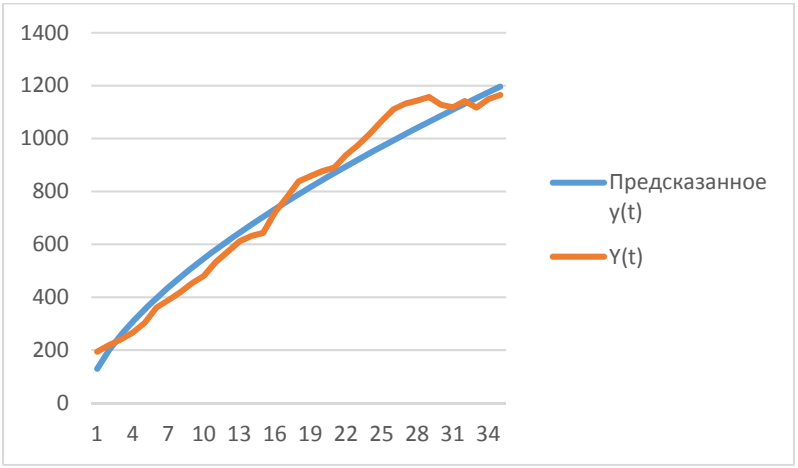


Рис. 26. График сравнения для нелинейной регрессии

7 этап. Интерпретация нелинейной модели и общее заключение

Построенная модель $y_t = 129{,}4191t^{0,6256}$ достаточно хорошо согласуется с имеющейся выборкой. Об этом свидетельствует значение индекса корреляции $\eta = 0{,}91553109$ и приемлемое значение средней ошибки аппроксимации $\bar{A} = 8\%$. Оценка регрессионной зависимости проводилась для значений t из промежутка от 1 до 35, поэтому построенная модель может быть использована для прогнозов среднего объясняемой переменной на этом промежутке и для значений t, близких к этому промежутку.

ПРИЛОЖЕНИЕ 4

Инструменты прогнозирования в Microsoft Excel

Прогнозирование – это очень важный элемент практически любой сферы деятельности, начиная от экономики и заканчивая инженерией. Существует большое количество программного обеспечения, специализирующегося именно на этом направлении. К сожалению, далеко не все пользователи знают, что обычный табличный процессор Excel имеет в своем арсенале инструменты для выполнения прогнозирования, которые по своей эффективности мало чем уступают профессиональным программам.

Целью любого прогнозирования является выявление текущей тенденции, и определение предполагаемого результата в отношении изучаемого объекта на определенный момент времени в будущем.

Процедура прогнозирования может быть выполнена следующими способами:

- Линия Тренда,
- Функция ПРЕДСКАЗ,
- Функция ТЕНДЕНЦИЯ,
- Функция РОСТ,
- Функция ЛИНЕЙН,
- Функция ЛГРФПРИБЛ.

Линия тренда

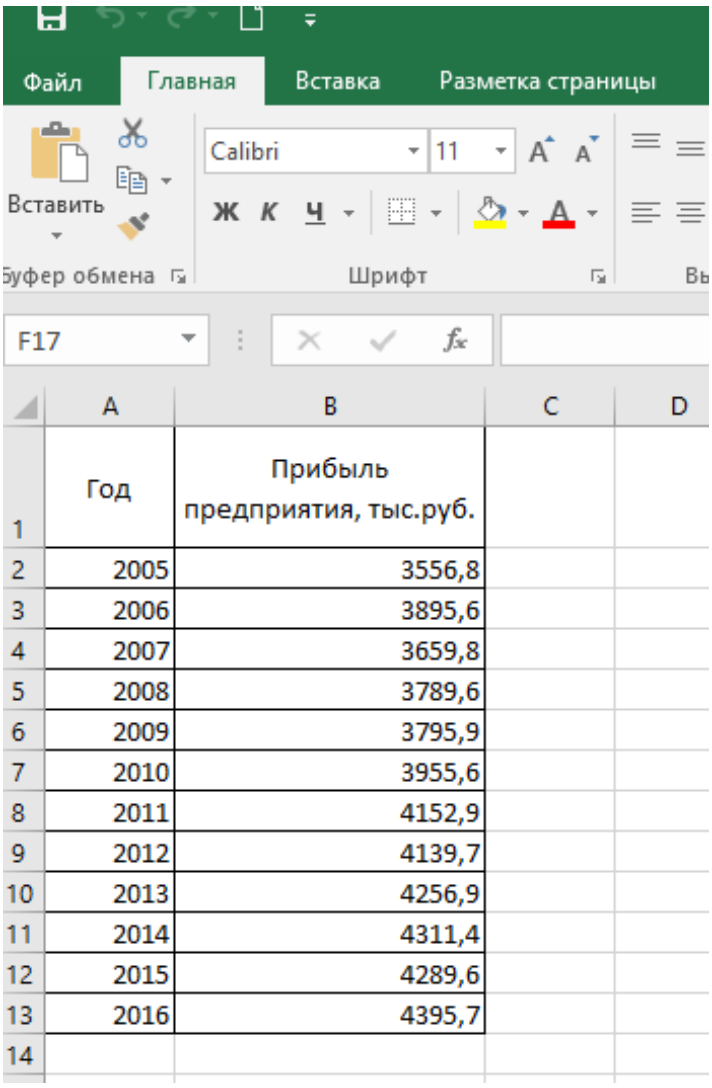
Одним из самых популярных видов графического прогнозирования в Excel является экстраполяция, выполненная построением линии тренда.

Пример 12. Рассмотрим временной ряд – сумма прибыли предприятия в тыс. руб. за 12 лет (рис. 27). Требуется предсказать сумму прибыли предприятия через 3 года.

1 шаг. Построим точечную диаграмму (рис. 28) для описания зависимости прибыли предприятия от года на основе табличных данных, состоящих из аргументов и значений функции.

2 шаг. Добавим линию тренда. Открывается окно форматирования линии тренда (рис. 29). В нем можно выбрать один из шести видов аппроксимации:

- Линейная;
- Логарифмическая;
- Экспоненциальная;
- Степенная;
- Полиномиальная;
- Линейная фильтрация.



	A	B	C	D
1	Год	Прибыль предприятия, тыс.руб.		
2	2005	3556,8		
3	2006	3895,6		
4	2007	3659,8		
5	2008	3789,6		
6	2009	3795,9		
7	2010	3955,6		
8	2011	4152,9		
9	2012	4139,7		
10	2013	4256,9		
11	2014	4311,4		
12	2015	4289,6		
13	2016	4395,7		
14				

Рис. 27. Исходные данные для тренда

Рис. 28. Точечная диаграммы для описания изменения прибыли предприятия

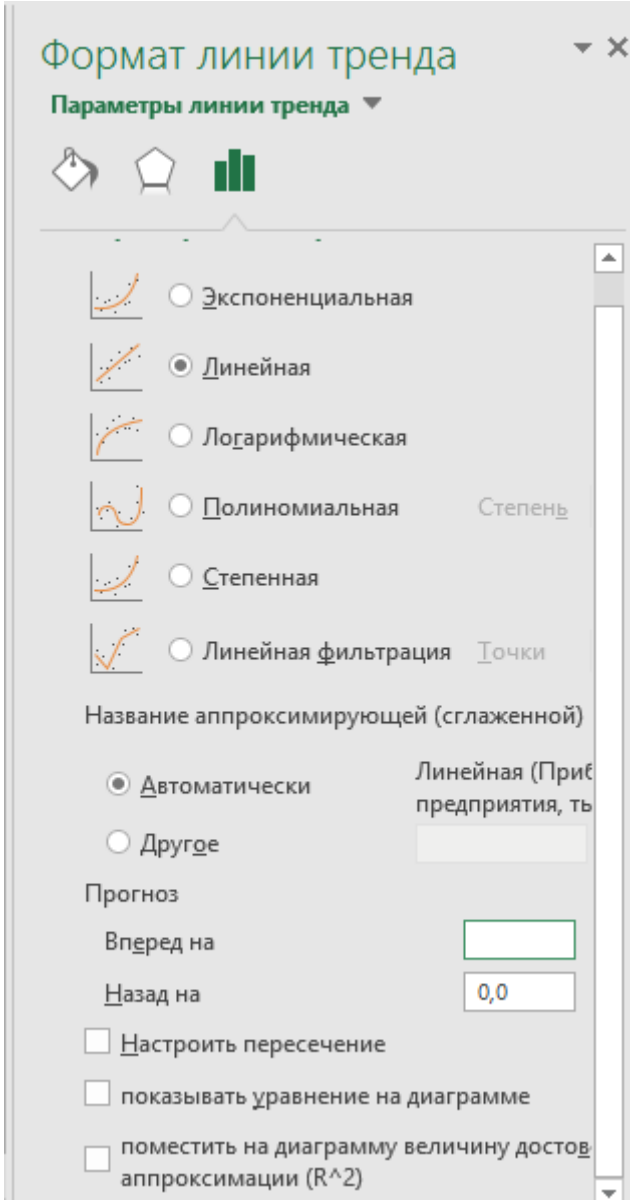


Рис. 29. Параметры форматирования тренда

3 шаг. Выберем линейную аппроксимацию. В блоке настроек «Прогноз» в поле «Вперед на» устанавливаем число «3,0», так как нам нужно составить прогноз на три года вперед. Кроме того, можно установить галочки около настроек «Показывать уравнение на диаграмме» и «Поместить на диаграмме величину достоверности аппроксимации (R^2)». Последний показатель отображает качество линии тренда, чем ближе его значение к 1, тем точнее уравнение тренда описывает зависимую переменную – сумму прибыли.

Рис. 30. Точечная диаграмма с трендом,
уравнением тренда и коэффициентом достоверности

Линия тренда построена и по ней мы можем определить примерную величину прибыли через три года. Как видим, к тому времени она должна перевалить за 4500 тыс. рублей. Коэффициент R^2, как уже было сказано выше, отображает качество линии тренда. В нашем случае величина R^2 составляет 0,89 (рис. 30). Чем выше коэффициент, тем выше достоверность линии. Максимальная величина его может быть равной 1. Принято считать, что при коэффициенте свыше 0,85 линия тренда является достоверной.

Если же уровень достоверности не устраивает, то можно вернуться в окно формата линии тренда и выбрать любой другой тип аппроксимации (рис. 31), или перепробовать все доступные варианты, чтобы найти наиболее точный.

Рис. 31. Точечная диаграмма с экспоненциальным трендом

Нужно заметить, что эффективным прогноз с помощью экстраполяции через линию тренда может быть, если период прогнозирования не превышает 30% от анализируемой базы периодов. То есть, при анализе периода в 12 лет мы не можем составить эффективный прогноз более чем на 3-4 года. Но даже в этом случае он будет относительно достоверным, если за это время не будет никаких форс-мажоров или наоборот чрезвычайно благоприятных обстоятельств, которых не было в предыдущих периодах.

Функция ПРЕДСКАЗ.ЛИНЕЙН

Экстраполяцию для табличных данных можно произвести через стандартную функцию табличного процессора Excel ПРЕДСКАЗ.ЛИНЕЙН. Она относится к категории статистических функций и имеет следующий синтаксис:

=ПРЕДСКАЗ.ЛИНЕЙН(X; известные_значения_y; известные значения_x)

«X» – это аргумент, значение функции, для которого нужно определить прогноз. В нашем случае в качестве аргумента будет выступать год, на который следует произвести прогнозирование.

«Известные значения y» — база известных значений функции. В нашем случае в её роли выступает величина прибыли за предыдущие периоды.

«Известные значения x» — это аргументы, которым соответствуют известные значения функции. В их роли у нас выступает нумерация годов, за которые была собрана информация о прибыли предыдущих лет.

Естественно, что в качестве аргумента не обязательно должен выступать временной отрезок. Например, им может являться температура, а значением функции может выступать уровень расширения воды при нагревании.

Пример 13. Рассмотрим временной ряд – сумма прибыли предприятия в тыс. руб. за 12 лет (рис. 27). Требуется предсказать сумму прибыли предприятия через 3 года.

1 шаг. Выделить любую незаполненную ячейку на листе, куда планируется выводить результат прогнозирования и вызвать функцию ПРЕДСКАЗ.ЛИНЕЙН категории «Статистические» (рис. 32)

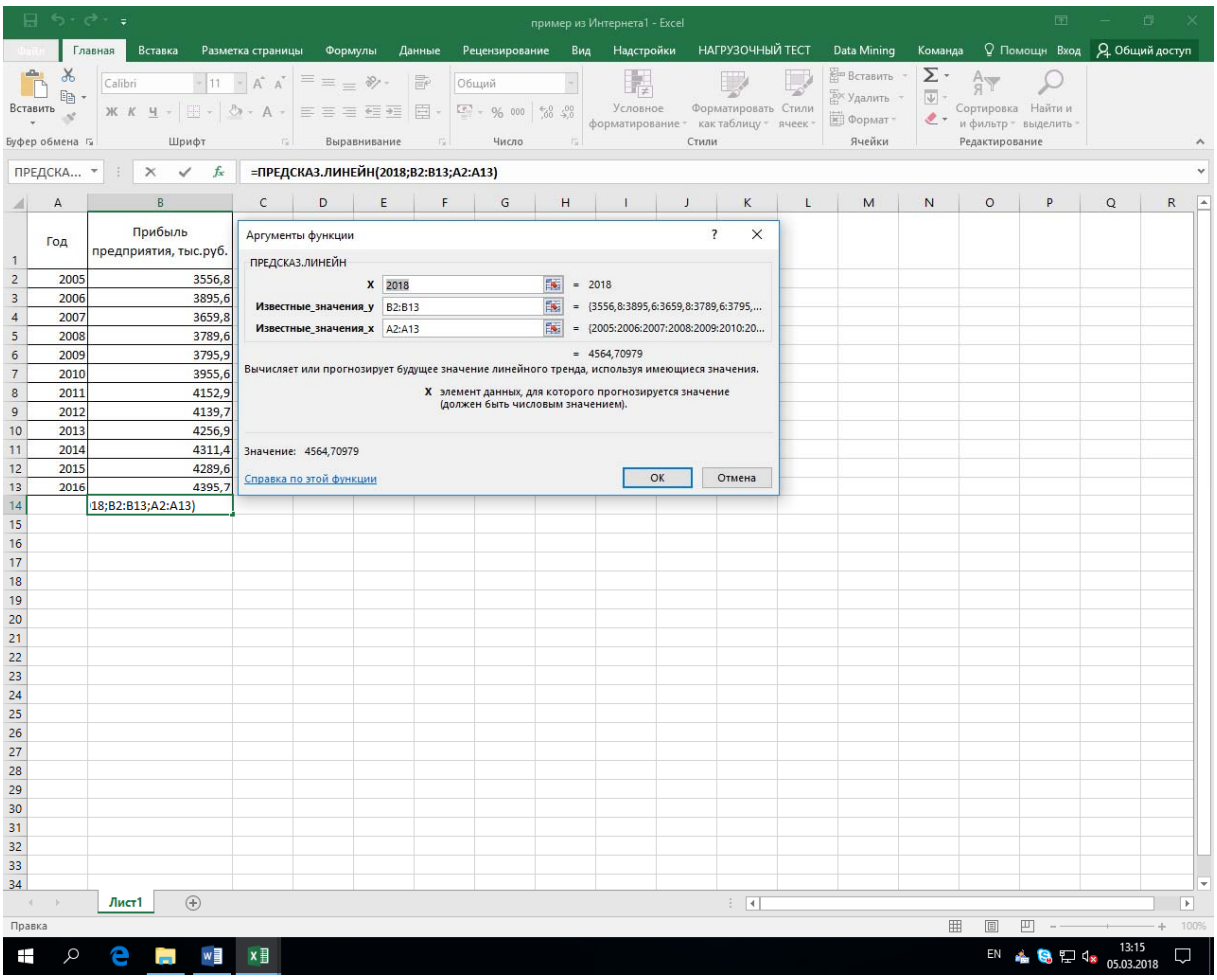


Рис. 32. Диалоговое окно функции ПРЕДСКАЗ.ЛИНЕЙН

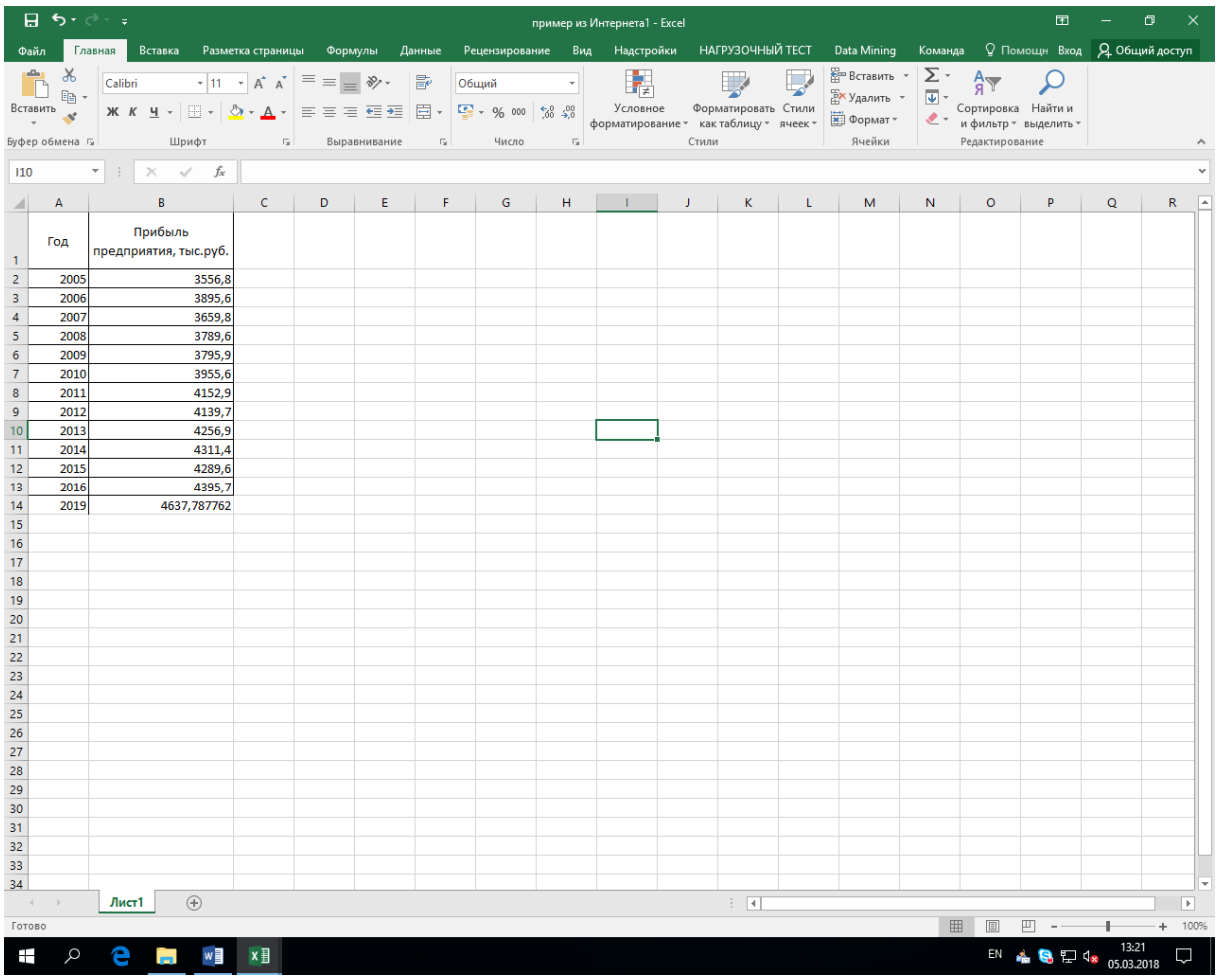
2 шаг. В поле «X» указывается величинА аргумента, к которому нужно отыскать значение функции. В нашем случаем это 2018 год. Поэтому вносим запись «2018». Но лучше указать этот показатель в ячейке на листе, а в поле «X» просто дать ссылку на него. Это позволит в будущем автоматизировать вычисления и при необходимости легко изменять год.

В поле «Известные значения y» указываются координаты столбца «Прибыль предприятия».

Аналогичным образом в поле «Известные значения x» вносится адрес столбца «Год» с данными за прошедший период.

На 2018 год планируется прибыль в районе 4564,7 тыс. рублей. На основе полученной таблицы можно построить график при помощи инструментов создания диаграммы.

Если поменять год в ячейке, которая использовалась для ввода аргумента, то соответственно изменится результат, а также автоматически обновится график. Например, по прогнозам в 2019 году сумма прибыли составит 4637,8 тыс. рублей (рис. 33).



Год	Прибыль предприятия, тыс.руб.
2005	3556,8
2006	3895,6
2007	3659,8
2008	3789,6
2009	3795,9
2010	3955,6
2011	4152,9
2012	4139,7
2013	4256,9
2014	4311,4
2015	4289,6
2016	4395,7
2019	4637,787762

Рис. 33. Прогноз на 2019 г.

Но не стоит забывать, что, как и при построении линии тренда, отрезок времени до прогнозируемого периода не должен превышать 30% от всего срока, за который накапливалась база данных.

Функция ТЕНДЕНЦИЯ

Для прогнозирования можно использовать ещё одну функцию – ТЕНДЕНЦИЯ. Она также относится к категории статистических операторов. Её синтаксис во многом напоминает синтаксис инструмента ПРЕДСКАЗ и выглядит следующим образом:

=ТЕНДЕНЦИЯ (Известные значения_y; известные значения_x; новые_значения_x; [конст])

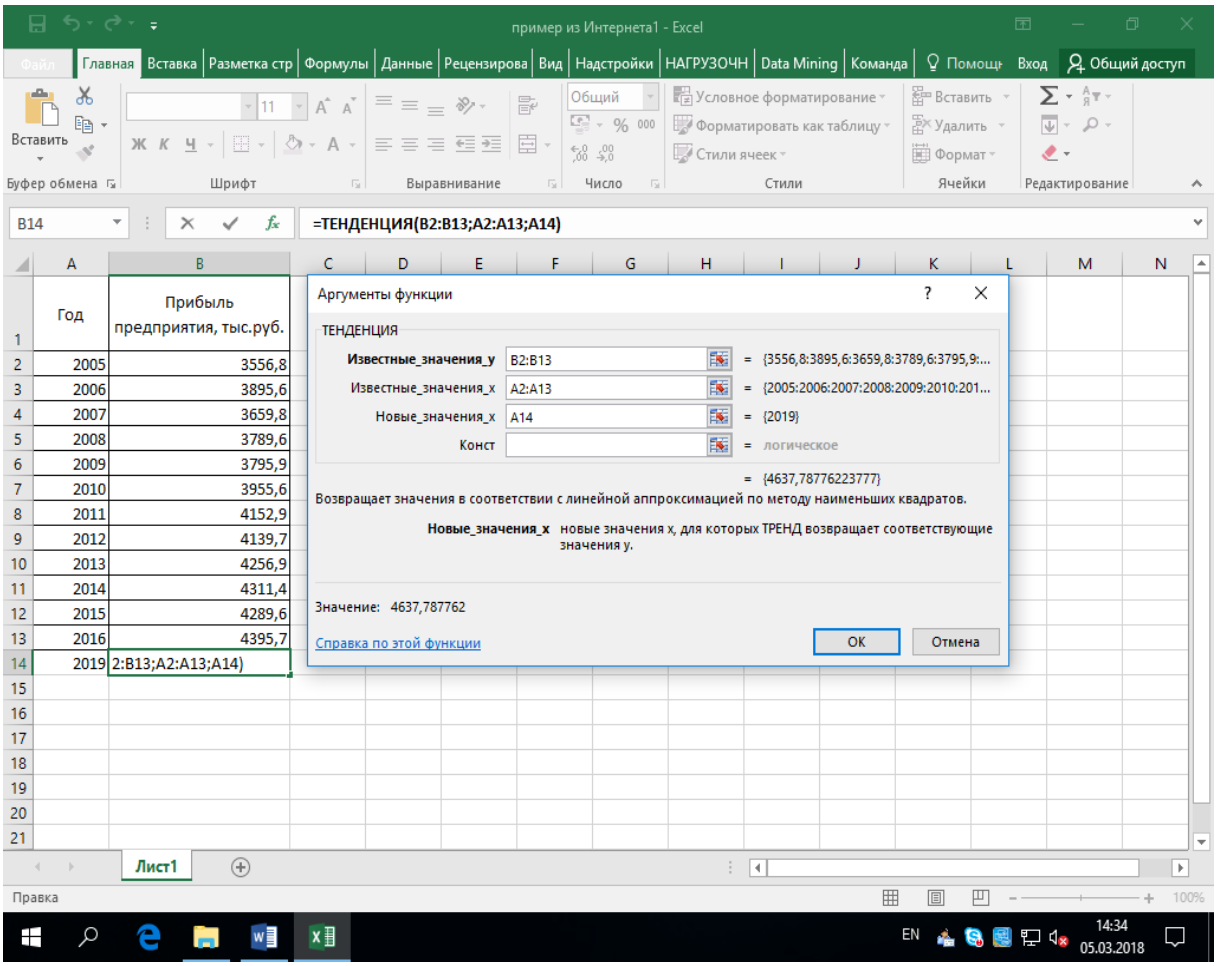
Как видим, аргументы «Известные значения y» и «Известные значения x» полностью соответствуют аналогичным элементам оператора ПРЕДСКАЗ.ЛИНЕЙН, а аргумент «Новые значения x» соответствует аргументу «X» предыдущего инструмента. Кроме того, у ТЕНДЕНЦИЯ имеется дополнительный аргумент «Константа», но он не является обязательным и используется только при наличии постоянных факторов.

Данный оператор наиболее эффективно используется при наличии линейной зависимости функции.

Пример 14. Рассмотрим временной ряд – сумма прибыли предприятия в тыс. руб. за 12 лет (рис. 27). Требуется предсказать сумму прибыли предприятия на 2019 г.

1 шаг. Установить курсор на ячейку для вывода результата и в категории «Статистические» выбрать функцию «ТЕНДЕНЦИЯ» (рис.34).

Рисунок 34.



Диалоговое окно функции ТЕНДЕНЦИЯ

2 шаг. Заполнить диалоговое окно. В поле «Известные значения y» занести координаты колонки «Прибыль предприятия». В поле «Известные значения x» ввести диапазон столбца «Год». В поле «Новые значения x» занести ссылку на ячейку, где находится номер года, для которого нужно вычислить прогноз – 2019 год. Поле «Константа» – пустое.

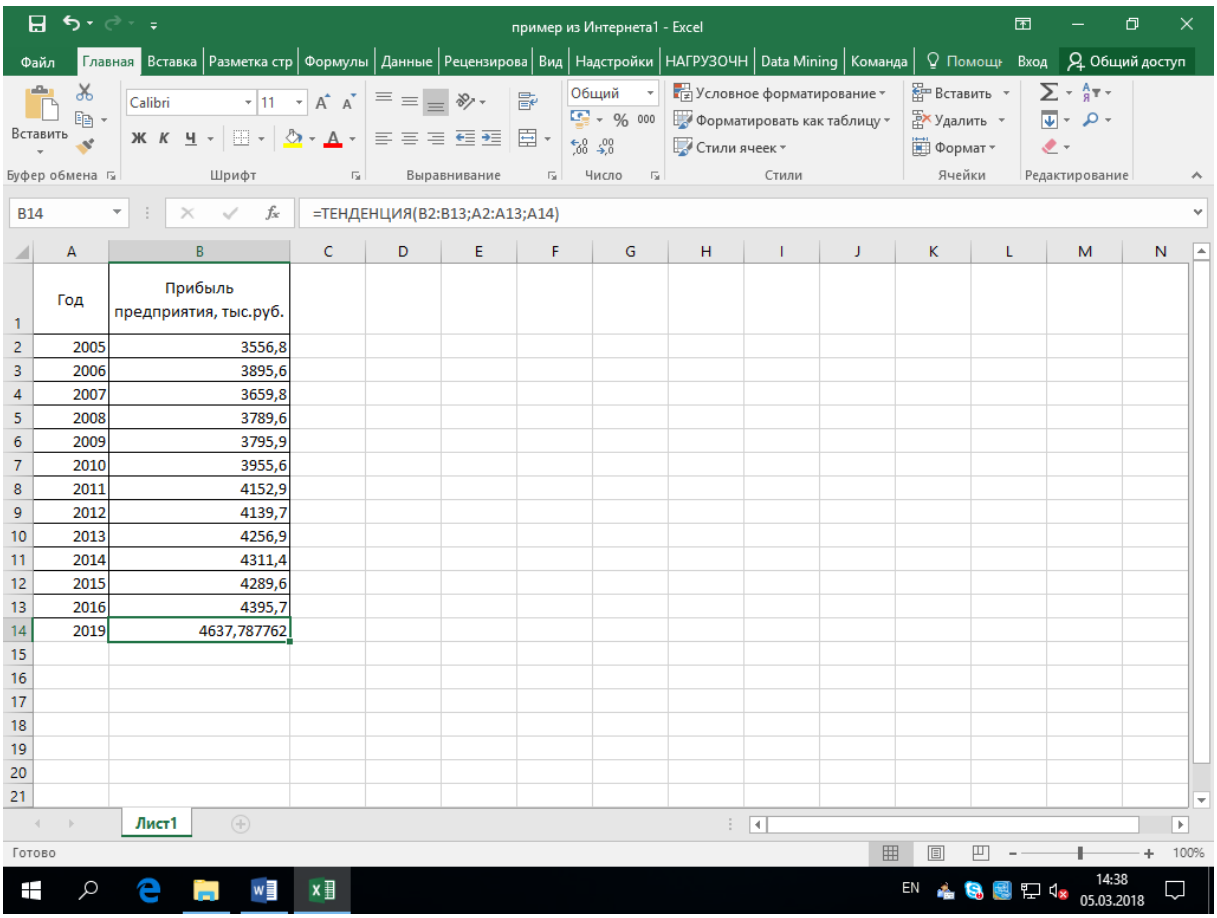


Рис. 35. Результат выполнения функции «ТЕНДЕНЦИЯ»

3 шаг. Функция обрабатывает данные и выводит результат на экран (Рис. 35). Как видим, сумма прогнозируемой прибыли на 2019 год, рассчитанная методом линейной зависимости, составит, как и при предыдущем методе расчета, 4637,8 тыс. рублей.

Функция РОСТ

Ещё одной функцией, с помощью которой можно производить прогнозирование в Excel, является оператор РОСТ. Он тоже относится к статистической группе инструментов, но, в отличие от предыдущих, при расчете применяет не метод линейной зависимости, а экспоненциальной. Синтаксис этого инструмента выглядит таким образом:

=РОСТ (Известные значения_y; известные значения_x; новые_значения_x; [конст])

Очевидно, что аргументы у данной функции в точности повторяют аргументы оператора ТЕНДЕНЦИЯ, так что второй раз на их описании останавливаться не будем, а сразу переходим к применению этого инструмента на практике.

Пример 15. Рассмотрим временной ряд – сумма прибыли предприятия в тыс. руб. за 12 лет (рис. 27). Требуется предсказать сумму прибыли предприятия на 2019 г.

1 шаг. Выделить ячейку вывода результата и в списке статистических операторов выбрать функцию «РОСТ» (рис. 36).

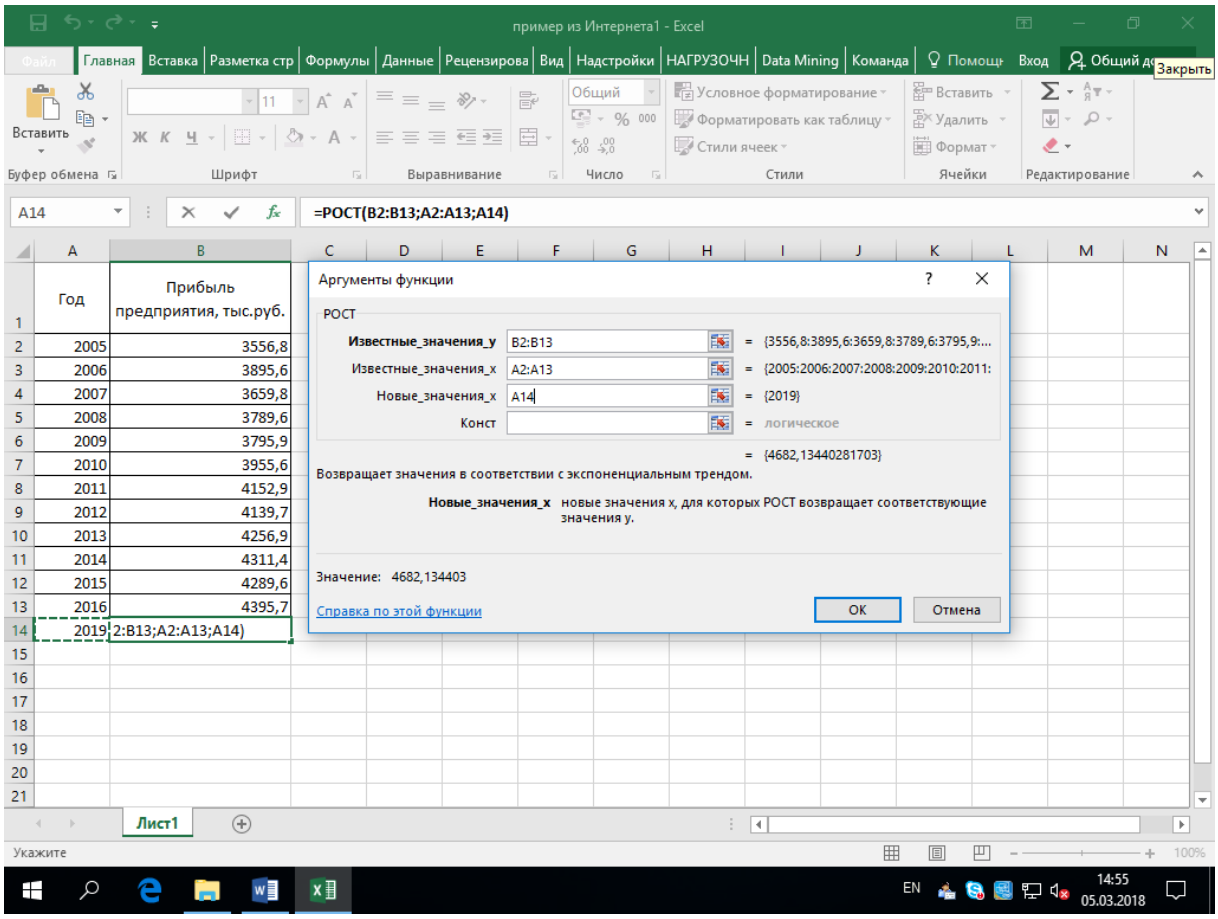


Рис. 36. Диалоговое окно функции «РОСТ»

2 шаг. Происходит активация окна аргументов указанной выше функции. Ввести в поля этого окна данные полностью аналогично тому, как они вводились в окне аргументов оператора ТЕНДЕНЦИЯ.

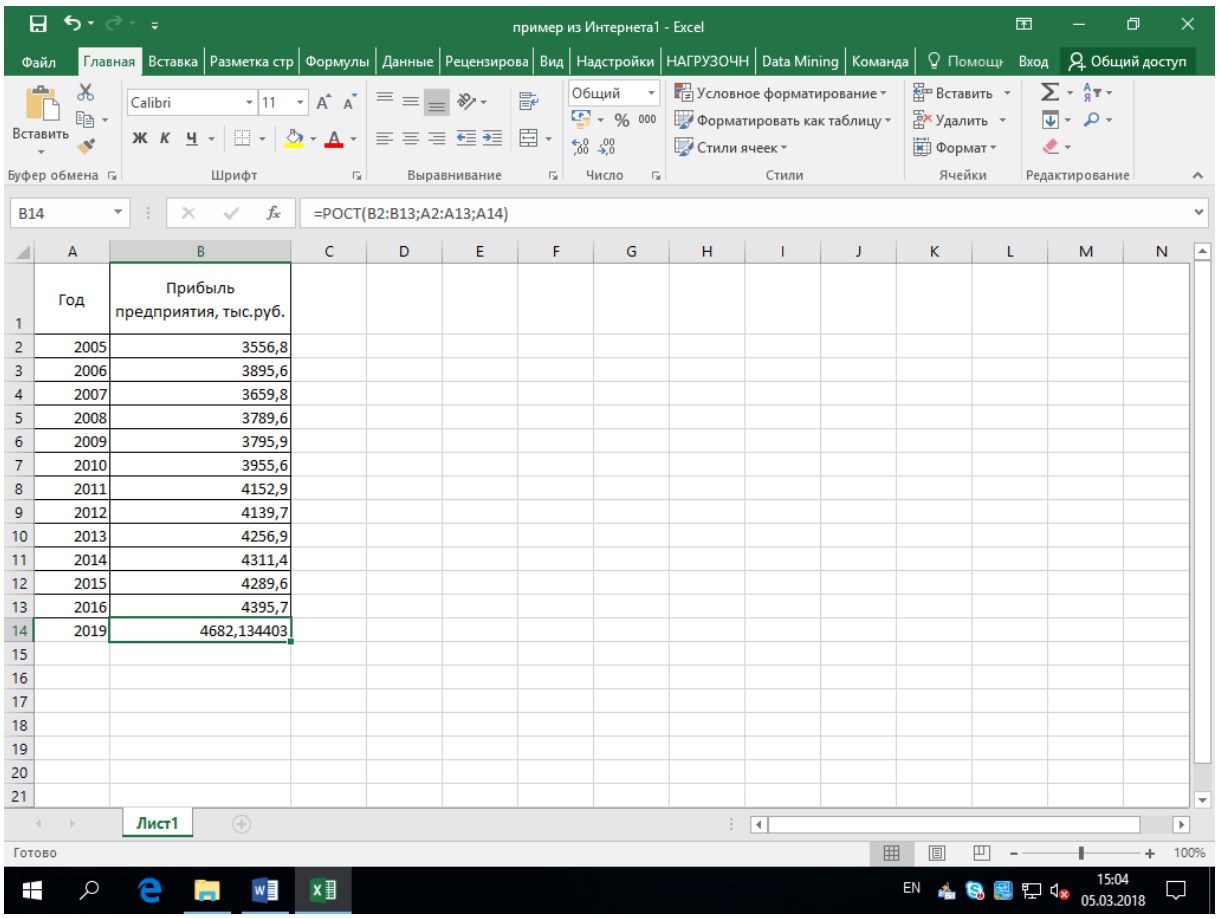


Рис. 37. Результат выполнения функции «РОСТ»

3 шаг. Результат обработки данных выводится на монитор в указанной ранее ячейке (рис. 37). Как видим, на этот раз результат составляет 4682,1 тыс. рублей. Отличия от результатов обработки данных оператором ТЕНДЕНЦИЯ незначительны, но они имеются. Это связано с тем, что данные инструменты применяют разные методы расчета: метод линейной зависимости и метод экспоненциальной зависимости.

Функция ЛИНЕЙН

Оператор ЛИНЕЙН при вычислении использует метод линейного приближения. Функция ЛИНЕЙН рассчитывает статистику для ряда с применением метода наименьших квадратов, чтобы вычислить прямую линию, которая наилучшим образом аппроксимирует имеющиеся данные и затем возвращает массив, который описывает полученную прямую. Функцию ЛИНЕЙН также можно объединять с другими функциями для вычисления других видов моделей, являющихся линейными по неизвестным параметрам, включая полиномиальные, логарифмические, экспоненциальные и степенные ряды. Поскольку возвращается массив значений, функция должна задаваться в виде формулы массива.

Уравнение для прямой линии имеет следующий вид: $y_t = ax_t + b$ или $y_t = b_0 + b_1x_{t1} + b_2x_{t2} + \ldots + b_kx_{tk}$

если существует несколько диапазонов значений x, где зависимые значения y — функции независимых значений x. Значения $b_1, \ldots, b_k$ — коэффициенты, соответствующие каждому значению x, а b_0 — постоянная. Обратите внимание, что y, x и b могут быть векторами. Функция ЛИНЕЙН возвращает массив $\{b_k, b_{k-1}, \ldots, b_0\}$. Функция ЛИНЕЙН может также возвращать дополнительную регрессионную статистику. Синтаксис этого инструмента выглядит таким образом:
ЛИНЕЙН(известные_значения_y; [известные_значения_x]; [конст]; [статистика])

- **Известные_значения_y.** Обязательный аргумент. Множество значений y, которые уже известны для соотношения $y_t = ax_t + b$.
  - Если массив ***известные_значения_y*** имеет один столбец, то каждый столбец массива ***известные_значения_x*** интерпретируется как отдельная переменная.
  - Если массив ***известные_значения_y*** имеет одну строку, то каждая строка массива ***известные_значения_x*** интерпретируется как отдельная переменная.
- **Известные_значения_x.** Необязательный аргумент. Множество значений x, которые уже известны для соотношения $y_t = ax_t + b$.
  - Массив ***известные_значения_x*** может содержать одно или несколько множеств переменных. Если используется только одна переменная, то массивы ***известные_значения_y*** и ***известные_значения_x*** могут иметь любую форму — при условии, что они имеют одинаковую размерность. Если используется более одной переменной, то ***известные_значения_y*** должны быть вектором (т. е. интервалом высотой в одну строку или шириной в один столбец).
  - Если массив ***известные_значения_x*** опущен, то предполагается, что это массив {1;2;3;...}, имеющий такой же размер, что и массив ***известные_значения_y***.
- **Конст.** Необязательный аргумент. Логическое значение, которое указывает, требуется ли, чтобы константа b была равна 0.
  - Если аргумент ***конст*** имеет значение ИСТИНА или опущен, то константа b вычисляется обычным образом.
  - Если аргумент ***конст*** имеет значение ЛОЖЬ, то значение b полагается равным 0 и значения a подбираются таким образом, чтобы выполнялось соотношение $y_t = ax_t$.
- **Статистика.** Необязательный аргумент. Логическое значение, которое указывает, требуется ли возвратить дополнительную регрессионную статистику.
  - Если аргумент ***статистика*** имеет значение ИСТИНА, функция **ЛИНЕЙН** возвращает дополнительную регрессионную статистику. Возвращаемый массив будет иметь следующий вид: **{mn;mn1;...;m1;b:sen;sen1;...;se1;seb:r2;sey:F;df:ssreg;ssresid}**.
  - Если аргумент ***статистика*** имеет значение ЛОЖЬ или опущен, функция **ЛИНЕЙН** возвращает только коэффициенты a и постоянную b.

Дополнительная регрессионная статистика выводится в виде:

Величина	**Описание**
se1,se2,...,sen	Стандартные значения ошибок для коэффициентов $b_1,...,b_k$.
seb	Стандартное значение ошибки для постоянной b_0 (seb = #Н/Д, если аргумент ***конст*** имеет значение ЛОЖЬ).
r2	Коэффициент детерминации. Сравниваются фактические значения y и значения, получаемые из уравнения прямой; по результатам сравнения вычисляется коэффициент детерминации, нормированный от 0 до 1. Если он равен 1, то имеет место полная корреляция с моделью, т. е. различий между фактическим и оценочным значениями y нет. В противоположном случае, если коэффициент детерминации равен 0, использовать уравнение регрессии для предсказания значений y не имеет смысла.
sey	Стандартная ошибка для оценки y.
F	F-статистика или F-наблюдаемое значение. F-статистика используется для определения того, является ли случайной наблюдаемая взаимосвязь между зависимой и независимой переменными.
df	Степени свободы. Степени свободы используются для нахождения F-критических значений в статистической таблице. Для определения уровня надежности модели необходимо сравнить значения в таблице с F-статистикой, возвращаемой функцией **ЛИНЕЙН**.
ssreg	Регрессионная сумма квадратов.
ssresid	Остаточная сумма квадратов.

Замечание

Если имеется только одна независимая переменная x, можно получить оценки параметров a и b для уравнения регрессии $y_t = ax_t + b$ непосредственно, воспользовавшись следующими формулами:

Оценка параметра a:

=ИНДЕКС(ЛИНЕЙН(известные_значения_y; известные_значения_x);1)

Оценка параметра b:

=ИНДЕКС(ЛИНЕЙН(известные_значения_y; известные_значения_x);2)

Пример 16. Рассмотрим временной ряд – сумма прибыли предприятия в тыс. руб. за 12 лет (рис. 26). Требуется оценить линейную модель парной регрессии и рассчитать прогноз прибыли на 2019 год.

1 шаг. Установить курсор на ячейку, начиная с которой будет производиться вычисление, выделить 10 ячеек: в две строки и пять столбцов, и вызвать Мастер функций – ЛИНЕЙН (рис. 38).

2 шаг. В поле «Известные значения y», открывшегося окна аргументов, вводим координаты столбца «Прибыль предприятия». В поле «Известные значения x» вносим адрес колонки «Год». Остальные поля заполняем единицами. Затем нажимаем комбинацию клавиш Ctrl–Shift–Enter (рис. 39).

Рассмотрим каждое полученное значение:

73,07797 – оценка параметра b,

-142906,6378 – оценка параметра *a*,
7,945366 – стандартная ошибка оценки параметра *b*,
15974,18106 – стандартная ошибка оценки параметра *a*,

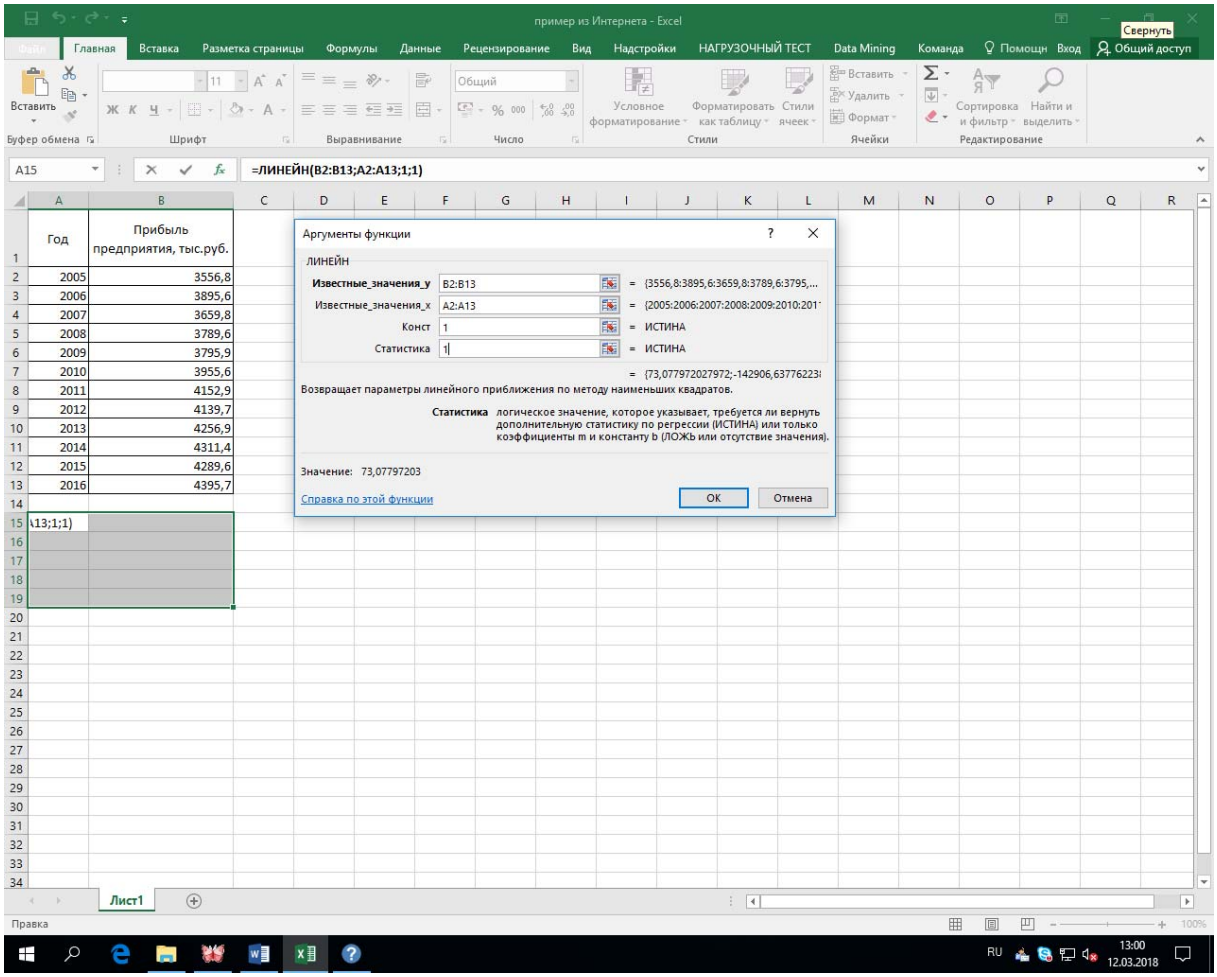
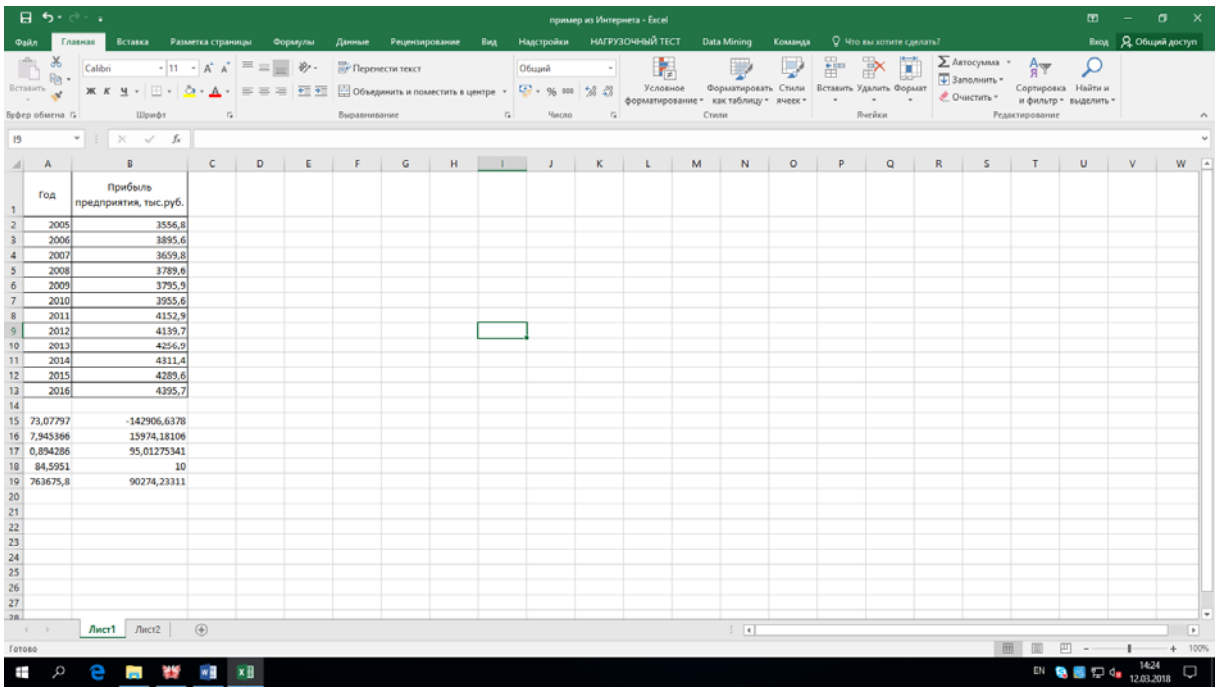


Рис. 38. Диалоговое окно функции ЛИНЕЙН

0,894286 – коэффициент детерминации,
95,01275341 – стандартная ошибка модели,
84,5951 – F – статистика,
10 – число степеней свободы,
763675,8 – объясненная регрессией часть дисперсии переменной Y,
90274,23311 – необъясненная регрессией часть дисперсии переменной Y.

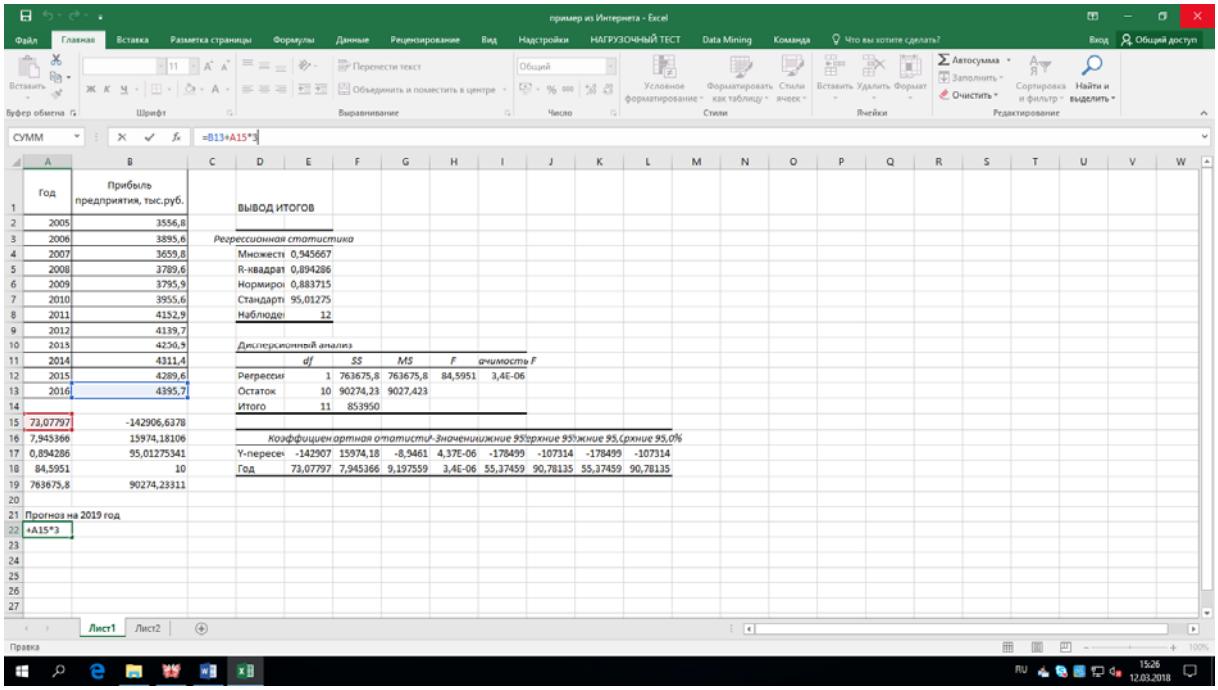
Если не выделять ячейки и ввести функцию в текущую ячейку, то на экране будет отображаться только одно значение оценки параметра *b*.



Год	Прибыль предприятия, тыс.руб.
2005	3556,8
2006	3895,6
2007	3659,8
2008	3789,6
2009	3795,9
2010	3955,6
2011	4152,9
2012	4139,7
2013	4256,9
2014	4311,4
2015	4289,6
2016	4395,7
73,07797	-142906,6378
7,945366	15974,18106
0,894286	95,01275341
84,5951	10
763675,8	90274,23311

Рис. 39. Результат выполнения вставки функции

Для сравнения рядом с результатами выполнения функции ЛИНЕЙН выведен диапазон ячеек с данными по регрессии (рис.40).



Год	Прибыль предприятия, тыс.руб.
2005	3556,8
2006	3895,6
2007	3659,8
2008	3789,6
2009	3795,9
2010	3955,6
2011	4152,9
2012	4139,7
2013	4256,9
2014	4311,4
2015	4289,6
2016	4395,7
73,07797	-142906,6378
7,945366	15974,18106
0,894286	95,01275341
84,5951	10
763675,8	90274,23311
Прогноз на 2019 год	
=A15*3	

ВЫВОД ИТОГОВ

Регрессионная статистика	
Множеств	0,945667
R-квадрат	0,894286
Нормиров	0,883715
Стандарт	95,01275
Наблюде	12

Дисперсионный анализ	df	SS	MS	F	значимость F
Регресси	1	763675,8	763675,8	84,5951	3,4E-06
Остаток	10	90274,23	9027,423		
Итого	11	853950			

	Коэффициенты	стандартная ошибка	t-статистика	P-значение	нижние 95%	верхние 95%	нижние 95,0%	верхние 95,0%
Y-пересеч	-142907	15974,18	-8,9461	4,37E-06	-178499	-107314	-178499	-107314
Год	73,07797	7,945366	9,197559	3,4E-06	55,37459	90,78135	55,37459	90,78135

Рис. 40. Сравнение результатов выполнения функции ЛИНЕЙН и регрессии

Теперь для расчета прогнозируемой прибыли на 2019 год ввести формулу в любую свободную ячейку, например, в A22:
= B13+A15*3,

где B13 – ячейка, в которой содержится фактическая величина прибыли за последний изучаемый год (2016 г.), а A15 – ячейка, в которой содержится оценка параметра *b* рассчитанного ранее линейного тренда, умножаем на 3 – так как между последним годом изучаемого периода (2016 г.) и годом на который нужно сделать прогноз (2019 г.) лежит срок в три года.

Как видим, прогнозируемая величина прибыли, рассчитанная методом линейного приближения, в 2019 году составит 4614,9 тыс. рублей (рис. 40).

Список иллюстраций

Варианты заданий для специальности 38.03.08

Бизнес-информатика.

Номер вашего варианта соответствует номеру студента в списке группы.

Вариант 1

Имеется временной ряд y_t, $t = \overline{1, 47}$, содержащий данные об объеме контейнерных перевозок грузов железнодорожным транспортом, млн. тонн (таблица 1). Требуется:

1 Выявить и исключить из y_t аномальные уровни ряда;
2 Определить наличие тренда в y_t методом серий и методом разности средних;
3 Сгладить временной ряд с помощью простой скользящей средней (период сглаживания m = 5) и экспоненциального сглаживания (параметр сглаживания γ = 0,4);
4 Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5 Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью p = 0,95;
6 Оценить качество полученных результатов;
7 Сформировать отчет (в электронном виде).

Вариант 2

Имеется временной ряд y_t, $t = \overline{1, 50}$, содержащий данные об объеме перевозок грузов железнодорожным транспортом, млн. тонн (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом разности средних;
3. Сгладить временной ряд с помощью взвешенной скользящей средней (период сглаживания m = 7) и экспоненциального сглаживания (параметр сглаживания γ = 0,1);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью p = 0,99;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 3

Имеется временной ряд y_t, $t = \overline{1, 46}$, содержащий данные об удельном весе автомобильного транспорта в объеме перевозок грузов транспортом общего пользования, % (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом Фостера-Стюарта;
3. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания m = 3) и экспоненциального сглаживания (параметр сглаживания γ = 0,7);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью p = 0,9;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 4

Имеется временной ряд y_t, $t = \overline{1, 48}$, содержащий данные о грузообороте морского вида транспорта общего пользования, млрд. тонно-км. (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом серий и методом разности средних;
3. Сгладить временной ряд с помощью взвешенной скользящей средней (период сглаживания m = 9) и экспоненциального сглаживания (параметр сглаживания γ = 0,6);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью p = 0,95;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 5

Имеется временной ряд y_t, $t = \overline{1, 50}$, содержащий данные об объеме грузов международного сообщения водным транспортом общего пользования, млн. тонн (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом серий и методом разности средних;

3. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания m = 5) и экспоненциального сглаживания (параметр сглаживания γ = 0,45);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью p = 0,99;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 6

Имеется временной ряд y_t, $t = \overline{1, 47}$, содержащий данные об объеме перевозок грузов железнодорожным транспортом, млн. тонн (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом Фостера-Стюарта;
3. Сгладить временной ряд с помощью взвешенной скользящей средней (период сглаживания m = 5) и экспоненциального сглаживания (параметр сглаживания γ = 0,8);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью p = 0,9;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 7

Имеется временной ряд y_t, $t = \overline{1, 48}$, содержащий данные о темпах роста отправления грузов авиационным видом транспорта общего пользования, % (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом серий и методом разности средних;
3. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания m = 2) и экспоненциального сглаживания (параметр сглаживания γ = 0,75);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью p = 0,99;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 8

Имеется временной ряд y_t, $t = \overline{1, 48}$, содержащий данные об удельном весе железнодорожного транспорта в общем грузообороте транспорта общего пользования, % (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом Фостер-Стюарта;
3. Сгладить временной ряд с помощью взвешенной скользящей средней (период сглаживания $m = 7$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}9$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}95$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 9

Имеется временной ряд y_t, $t = \overline{1, 49}$, содержащий данные о затратах на перевозку пассажиров и грузов ФЖДТ (в относительных единицах по отношению к уровню 1991 г., таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом Фостера-Стюарта;
3. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания $m = 3$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}8$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}99$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 10

Имеется временной ряд y_t, $t = \overline{1, 48}$, содержащий данные о динамике изменения числа аэропортов в России, шт. (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом разности средних;

3. Сгладить временной ряд с помощью взвешенной скользящей средней (период сглаживания $m = 9$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}2$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}9$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 11

Имеется временной ряд y_t, $t = \overline{1, 46}$, содержащий данные о перевозках грузов железнодорожным транспортом, млн.т. (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом серий и методом Фостера-Стюарта;
3. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания $m = 3$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}35$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}95$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 12

Имеется временной ряд y_t, $t = \overline{1, 48}$, содержащий данные о производстве легковых автомобилей, тыс. шт. (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом разности средних;
3. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания $m = 2$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}85$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}95$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 13

Имеется временной ряд y_t, $t = \overline{1, 47}$, содержащий данные о производстве нефти, млн. т. (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом серий и методом Фостера-Стюарта;
3. Сгладить временной ряд с помощью взвешенной скользящей средней (период сглаживания $m = 5$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}9$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}9$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 14

Имеется временной ряд y_t, $t = \overline{1, 48}$, содержащий данные о производстве газа естественного, млрд. куб. м. (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом Фостера-Стюарта;
3. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания $m = 3$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}5$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}99$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 15

Имеется временной ряд y_t, $t = \overline{1, 50}$, содержащий данные об экспорте нефти из стран ОПЕК в развитые страны, млн. т. (таблица 1). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом Фостера-Стюарта;
3. Сгладить временной ряд с помощью взвешенной скользящей средней (период сглаживания $m = 5$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}85$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;

5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}99$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Таблица 1

Исходные данные

№ наблюдения	Номер варианта														
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1	23	58	10	810	39	40	120	88	14	3019	88,1	58,5	25,6	50,4	261,54
2	22	65	13	2500	22	25	200	87	18	2814	90,5	67	23,4	51,2	243,50
3	19	55	19	2100	29	17	245	90	30	2773	92	80,5	25,4	53,4	254,87
4	20	65	24	3700	18	35	230	74	60	2648	93,7	70,2	24,3	48,3	260,98
5	18	68	22	4500	17	49	302	80	101	2572	97,8	68,2	25,5	48,9	259,12
6	22	70	22	4190	26	22	284	90	139	2431	94,1	72,8	24,6	43,5	253,64
7	23	69	26	5080	11	17	370	85	284	2379	96,9	70,2	25,4	44	258,47
8	25	74	23	4930	24	20	290	81	346	2306	98,2	75,8	25,3	47,6	264,26
9	18	68	19	5275	13	12	301	82	473	2583	99,8	67,4	24,3	43,2	247,80
10	19	72	17	14200	29	33	277	88	586	2197	100,5	77,1	25,2	50,8	218,93
11	20	75	20	8100	27	31	324	85	491	2064	126,5	66,7	24,1	52,8	205,70
12	28	70	22	6900	28	28	320	87	601	2132	131,6	61,8	24,9	51	204,96
13	30	77	18	9150	14	39	321	83	769	2075	138	65,4	23,1	51,7	198,80
14	26	79	23	9280	27	42	348	86	637	2057	144,8	69,9	20,9	55,2	174,86
15	29	72	29	12900	20	35	360	90	821	2031	159,1	74,3	19,4	58,7	163,08
16	34	82	27	10965	24	49	327	95	932	1987	171	75,6	19,3	53,1	171,61
17	38	80	29	11295	28	57	305	78	843	1964	185,4	79,8	19,6	58,6	160,47
18	45	74	25	10500	48	76	293	73	1025	1935	193,3	82,9	18,9	59,9	134,19
19	53	79	31	12365	38	90	300	60	996	1872	212,2	81,2	18,3	60,6	139,46
20	69	80	33	14280	45	66	280	81	1074	1724	235,3	83,4	17,4	61,2	188,43
21	70	81	35	13795	57	79	237	85	1072	1705	243,5	81,4	17,2	59,8	116,31
22	81	87	46	18975	46	100	229	93	1161	1697	256,7	93,5	18,2	63,2	110,26
23	89	84	32	12850	71	79	208	89	1095	1683	274,3	94,4	16,7	55,7	131,33
24	90	90	46	18295	77	110	215	92	987	1678	295,8	91,2	15,6	64,8	136,58
25	85	85	39	20400	79	140	236	95	1243	1671	321,5	85,6	15,2	63,4	126,54
26	86	84	37	25900	85	109	228	99	1184	1664	365,7	89,9	15,9	60,1	126,04
27	92	93	43	15845	99	145	215	84	1266	1642	388,9	93,4	14,7	69,7	119,17
28	84	95	50	22950	120	136	239	79	1201	1625	413,2	99,8	13,9	75,4	113,67
29	75	90	40	28635	95	124	284	81	1375	1609	423,3	102,2	12,4	76,5	113,79
30	69	99	48	23480	86	150	231	90	1292	1588	421,2	99,7	13,4	78,6	104,69
31	72	96	48	20598	149	110	224	96	1349	1523	436,5	100,6	13,5	79,6	102,70
32	69	101	49	25500	174	101	217	88	1157	1511	513,5	100,9	12,3	70,9	123,43
33	55	94	46	30850	160	129	219	85	1254	1374	535,6	100,9	12,9	79,7	126,02
34	59	95	47	23990	155	116	235	74	1373	1332	687,8	99,8	10,2	84,3	130,77
35	40	100	38	28635	168	103	228	80	1486	1275	689,9	103,7	10,9	85,1	145,90
36	33	103	47	34925	210	99	220	81	1398	1176	715,2	120,3	11,5	90,1	127,29
37	39	101	48	27800	200	114	215	86	1461	1028	735,6	123,5	12,4	89,7	121,55
38	26	98	54	30210	240	92	210	89	1525	1012	812,6	113,5	11,6	91,2	117,26
39	29	100	56	40935	249	68	212	94	1587	994	896,8	129,8	10,8	90,3	136,07
40	22	110	56	22840	227	51	219	93	1614	986	924,3	135,6	11,2	93,6	141,10
41	34	109	56	36000	294	13	205	91	1683	971	967,8	130,3	11,8	98,6	129,01
42	28	98	57	30860	261	40	194	86	1736	965	1003,5	142,3	10,9	99,1	127,09
43	20	105	59	29100	326	49	183	87	1790	943	1346,8	130,7	10,4	96,5	136,34
44	27	104	54	40285	308	23	168	83	1921	937	1653,8	145,6	12,9	98,7	149,70

№ наблюдения	Номер варианта														
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
45	11	117	55	40055	357	38	159	90	1868	912	1890,9	159,8	10,9	99,2	131,32
46	28	108	57	35800	384	19	127	86	1837	898	2003,1	176,4	10,3	93,6	140,77
47	14	120		46895	320	22	101	89	1974	884		198,5	10,6	89,7	158,94
48		127		38700	363		119	94	2012	871		232,6		97,8	182,70
49		113			450				2106						167,27
50		116			425										161,94

Варианты заданий для специальности 09.03.03

Прикладная информатика.

Номер вашего варианта соответствует номеру в списке группы.

Вариант 1

Имеется временной ряд y_t, $t = \overline{1, 49}$, содержащий данные об индексе потребительских цен (в % к предыдущему месяцу) платных услуг (таблица 2). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом серий и методом разности средних;
3. Сгладить временной ряд с помощью взвешенной скользящей средней (период сглаживания $m = 7$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}55$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}99$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 2

Имеется временной ряд y_t, $t = \overline{1, 49}$, содержащий данные об объемах продаж, млн. руб. (таблица 2). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом серий и методом разности средних;
3. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания $m = 2$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}35$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}95$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 3

Имеется временной ряд y_t, $t = \overline{1, 49}$, содержащий данные о курсе акций компании IBM, долл. (таблица 2). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом серий и методом Фостера-Стюарта;
3. Сгладить временной ряд с помощью взвешенной скользящей средней (период сглаживания $m = 7$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}7$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}99$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 4

Имеется временной ряд y_t, $t = \overline{1, 47}$, содержащий данные о потенциальной платежеспособности рынка по тарифному плану «Таксофон», млн. руб. (таблица 2). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом разности средних;
3. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания $m = 6$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}3$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}9$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 5

Имеется временной ряд y_t, $t = \overline{1, 48}$, содержащий данные о затратах на обновление бланков АО “Электросвязь”, тыс. руб. (таблица 2). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом Фостера-Стюарта;
3. Сгладить временной ряд с помощью взвешенной скользящей средней (период сглаживания $m = 7$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}45$);

4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}99$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 6

Имеется временной ряд $y_t, \quad t = \overline{1, 47}$, содержащий данные о продаже сотовых телефонов Nokia, млн. руб. (таблица 2). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом серий и методом Фостера-Стюарта;
3. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания $m = 5$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}8$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}9$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 7

Имеется временной ряд $y_t, \quad t = \overline{1, 45}$, содержащий данные о динамике количества пользователей сети “Радиотел”, тыс. чел. (таблица 2). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом серий и методом разности средних;
3. Сгладить временной ряд с помощью взвешенной скользящей средней (период сглаживания $m = 5$) и экспоненциального сглаживания (параметр сглаживания $\gamma = 0{,}5$);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}95$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 8

Имеется временной ряд y_t, $t = \overline{1, 50}$, содержащий данные о накоплении капитала компаниями связи по России, млн. долл. (таблица 2). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом разности средних;
3. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания m = 4) и экспоненциального сглаживания (параметр сглаживания γ = 0,2);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью p = 0,99;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 9

Имеется временной ряд y_t, $t = \overline{1, 50}$, содержащий данные о реализации таксофонных карт, тыс. шт. (таблица 2). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом “восходящих” и “нисходящих” серий и методом Фостера-Стюарта;
3. Сгладить временной ряд с помощью взвешенной скользящей средней (период сглаживания m = 9) и экспоненциального сглаживания (параметр сглаживания γ = 0,1);
4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью p = 0,9;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Вариант 10

Имеется временной ряд y_t, $t = \overline{1, 49}$, содержащий данные об объеме финансирования в отрасль связи, млн. руб. (таблица 2). Требуется:

1. Выявить и исключить из y_t аномальные уровни ряда;
2. Определить наличие тренда в y_t методом серий и методом разности средних;
3. Сгладить временной ряд с помощью простой скользящей средней (период сглаживания m = 7) и экспоненциального сглаживания (параметр сглаживания γ = 0,75);

4. Обосновать выбор уравнения тренда и оценить его параметры с помощью метода наименьших квадратов;
5. Построить точечный и интервальный прогнозы на 4 шага вперед с доверительной вероятностью $p = 0{,}95$;
6. Оценить качество полученных результатов;
7. Сформировать отчет (в электронном виде).

Таблица 2

Исходные данные

№ наблюдения	Номер варианта									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	123	200	510	1000	650	110,11	37	4200	400	22
2	119	310	497	1090	860	104,24	29	4300	310	24
3	110	320	504	1050	1000	100,39	22	5500	330	26
4	112	260	510	1030	1250	101,56	25	7000	290	26
5	111	190	509	1010	1200	100,75	15	7400	210	32
6	109	210	503	1040	1247	100,96	19	8200	200	27
7	110	310	500	1020	1350	104	18	8500	220	33
8	108	410	500	1025	1460	102,14	20	9800	180	35
9	107	430	500	1045	1450	102	17	10650	110	39
10	109	370	495	900	1300	101,71	21	12050	70	37
11	107	300	494	890	1500	103	13	11900	120	41
12	103	320	499	1040	1615	100,64	18	12400	50	43
13	100	340	489	999	1650	103	15	11500	35	41
14	102	350	483	850	1700	103,36	16	13200	70	47
15	103	410	483	990	1700	104,75	17	13970	90	49
16	99	430	497	1037	1800	104	18	13300	20	45
17	98	450	473	1078	1650	104,9	14	13900	15	50
18	102	390	469	1100	1750	106	15	14000	10	50
19	94	410	470	1081	1800	105,9	18	14100	5	50
20	87	450	478	989	1750	106,8	12	14700	9	51
21	89	470	448	945	1810	106,1	15	14870	7	50
22	100	490	432	958	1890	109,1	17	15000	8	56
23	86	520	439	899	1800	107,6	19	15200	20	54
24	84	490	432	911	2000	108,3	17	15800	30	57
25	85	530	421	945	1600	110	15	14830	20	57
26	83	550	389	993	1900	107	19	14270	35	59
27	84	610	345	980	1950	110	17	13000	50	55
28	82	650	323	995	1930	110,64	14	15050	70	61
29	79	590	340	1052	1950	112,5	16	16000	60	61
30	74	610	321	1101	1960	111,31	13	14370	100	62
31	76	630	312	919	1975	111,71	16	14030	130	61
32	73	670	300	948	2000	114,3	14	13900	110	65
33	74	710	248	987	2010	112	14	13800	190	62
34	69	740	215	956	2015	114,96	15	11000	210	68
35	70	780	236	1049	2020	115,1	17	13050	250	69
36	65	790	230	1112	2030	117,5	12	11200	290	70
37	63	810	241	931	2040	114	18	12050	230	70
38	67	790	215	1053	2050	118,24	19	10000	320	71
39	62	820	221	1031	2060	118	13	10500	390	71
40	57	850	254	931	2070	121	14	9900	340	71
41	59	910	239	943	2080	119	15	9000	400	78
42	53	890	217	1046	2090	121,5	12	7500	410	80
43	48	870	201	976	2100	122,54	14	7900	500	82

№ наблюдения	Номер варианта									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
44	50	860	212	1076	2100	121,82	18	7750	485	91
45	51	910	231	903	2115	124,2	17	5100	550	93
46	46	930	203	1067	2125	127		4320	510	97
47	42	910	219	915	2135	123,84		4080	685	100
48	43	930	201		2140			4100	720	98
49	41	920	195					2000	880	97
50								900	840	

Учебное издание

Белых Татьяна Ивановна, **Бурдуковская** Анна Валерьевна

ПРЕДВАРИТЕЛЬНЫЙ АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

Учебное пособие

Издается в авторской редакции

ИД № 06318 от 26.11.01.
Подписано в пользование 28.05.18

Издательство Байкальского государственного университета.
664003, г. Иркутск, ул. Ленина, 11.
htpp://bgu.ru.