

*Рабочий вариант
Исправлено 14.09.2017
Кроме Приложения 3*

Н.В. Карпенко

**Эконометрика
Часть 3**

**Анализ и прогнозирование
временного ряда**

Учебное пособие

Оглавление

Введение

1. Понятия временных рядов
 - 1.1. Этапы моделирование динамики социально-экономических процессов
 - 1.2. Виды и классификация временных рядов
 - 1.3. Модели временных рядов
2. Общий анализ временного ряда
 - 2.1. Графический анализ временного ряда
 - 2.2. Проверка гипотезы о случайности тренда
 - 2.3 Численное сглаживание уровней временного ряда. Метод скользящей средней
 - 2.4. Исследование структуры ряда. Автокорреляция
3. Моделирование временного ряда без учета сезонности
 - 3.1. Виды уравнений тренда
 - 3.2. Построение линейного тренда
 - 3.3. Оценка качества уравнения линейного тренда
 - 3.4. Анализ отклонений от тренда
 - 3.5. Модель ряда и ее характеристики
 - 3.6. Моделирование временного ряда в условиях смены тенденции. Тест Чоу
4. Моделирование временного ряда с учетом сезонности
 - 4.1. Аддитивная модель с учетом сезонности
 - 4.2. Определение степени полиномиального тренда

- 5. Методы прогнозирования временных рядов
 - 5.1. Прогнозирование по среднему приросту
 - 5.2. Прогнозирование по среднему темпу роста
 - 5.3. Прогнозирование методом численного сглаживания
 - 5.4. Прогнозирование по уравнению тренда
 - 5.5. Прогнозирование по аддитивной модели с учетом сезонности

Приложения

Приложение 1

Приложение 2

Приложение 3

Приложение 4

Литература

Введение

В настоящее время в экономическом анализе актуальным является не только сравнение объектов между собой по какой-то группе показателей, но и сопоставление одного или нескольких показателей, характеризующих состояние одного объекта (процесса) за ряд последовательных периодов времени. Информационной базой для такого анализа служат динамические ряды.

Последовательность наблюдения одного явления или показателя, упорядоченного в зависимости от последовательно возрастающих или убывающих значений другого признака, называют **динамическим** рядом. Например, зависимость прибыли банка от объема депозитных вкладов.

Если в качестве признака, по которому производится упорядочивание, берется время, то такие ряды называют **временными** рядами. Данное учебное пособие посвящено изучению временных рядов.

Существует две основные цели анализа временных рядов: **определение природы ряда** и **прогнозирование**, т.е. предсказание будущих значений временного ряда по настоящим и прошлым значениям. Обе цели требуют, чтобы модель ряда была определена и более или менее формально описана. Как только модель построена, с ее помощью можно интерпретировать рассматриваемые данные - например, использовать ее для анализа сезонного изменения цен на товары. Затем можно экстраполировать ряд на основе найденной модели, т.е. предсказать его будущие значения.

В настоящем учебном пособии изложены основные этапы общего анализа временного ряда, методы построения и оценки качества моделей временного ряда, методы прогнозирования и оценки точности прогноза.

Практическое приложение теоретических положений показано на сквозном примере эконометрического анализа доходов населения.

В учебном пособии даны методические рекомендации для проведения расчетов средствами Excel MS Office.

1. Понятия временных рядов

1.1. Этапы моделирование динамики социально-экономических процессов

Важную роль в экономике играет анализ, моделирование и прогнозирование на базе данных по одному объекту, собранных в последовательные моменты или периоды времени, т.е. на основе временных рядов.

Первым этапом такого исследования является построение ряда, отвечающего требованиям статистической науки: уровни ряда должны быть получены из надежных источников информации и сопоставимы друг с другом по качественному содержанию.

Второй этап – определение типа основной тенденции динамики (либо констатация отсутствия надежной тенденции). Методика этого этапа излагается в курсах общей теории статистики и обычно включает:

– содержательный подход – на основе закономерностей экономики, технологии, общего состояния изучаемого объекта;

– графическое изображение временного ряда и подбор подходящей линии тренда;

– математико–статистическую оценку наличия достаточно надежного прироста уровней, ускорения этого прироста, темпов роста, наличия автокорреляции уровней.

Третий этап – вычисление уравнения тренда, т.е. уравнения такой линии, которая оптимально выражает фактическую тенденцию изменения уровней ряда. На этом этапе применяются методы математической статистики: метод наименьших квадратов (МНК), критерии оценки надежности – Фишера, Стьюдента, Дарбина–Уотсона и др. Используются программы STATISTICA, Statgraphics, средства Excel MS Office и др.

Четвертый этап заключается в исследовании отклонений фактических значений уровней ряда от расчетных уровней тренда, т.е. изучении колеблемости, а также в измерении и моделировании сезонных (циклических) колебаний.

Пятым этапом является расчет прогнозируемых значений временного ряда для будущих периодов, вероятных интервалов этих прогнозов, а иногда и страхового запаса.

1.2. Виды и классификация временных рядов

Развитие экономического явления или процесса во времени в статистике называется динамикой. Для отображения динамики строятся временные ряды, которые представляют собой ряды изменяющихся во времени значений статистического показателя, расположенных в хронологическом порядке.

Основной характеристикой временного ряда является его длина n . Под длиной временного ряда обычно понимается период времени, прошедший от первого наблюдения до последнего. Иногда под длиной временного ряда также понимают количество его элементов.

Элементами (членами) временного ряда являются показатели уровней ряда и периоды (годы, кварталы, месяцы, сутки) или моменты (даты) времени. Уровни ряда обычно обозначаются Y_t , а соответствующие им моменты времени - t .

Во временных рядах последовательные наблюдения (уровни), как правило, зависят друг от друга.

Информационная ценность уровней временного ряда уменьшается по мере их удаления от текущего момента времени.

Точность характеристик временного ряда зависит от числа наблюдений во временном ряду, но эта зависимость не является прямо пропорциональной.

Классифицировать различные виды временных рядов можно по следующим признакам.

1. В зависимости от способа выражения уровней временные ряды подразделяются на **ряды абсолютных, относительных и средних величин**.

2. В зависимости от того, как выражают уровни ряда состояние экономического явления на определенные моменты времени (начало месяца, квартала, года и т.п.) или его величину за определенные интервалы времени (например, за месяц, квартал, год и т.п.) различают соответственно **моментные и интервальные временные ряды**.

3. В зависимости от расстояния между уровнями временные ряды разделяются на **ряды с равноотстоящими уровнями и неравноотстоящими уровнями** во времени.

4. В зависимости от наличия основной тенденции изучаемого процесса временные ряды разделяются на **стационарные и нестационарные**.

Если основные характеристики случайного процесса (математическое ожидание и дисперсия) постоянны и не зависят от времени, то процесс считается стационарным и временной ряд также считается стационарным. График стационарного ряда располагается относительно прямой $y = y_{\text{ср}}$, параллельной горизонтальной оси. На рис. 1 среднее значение стационарного ряда $\bar{y}_t = 4$.

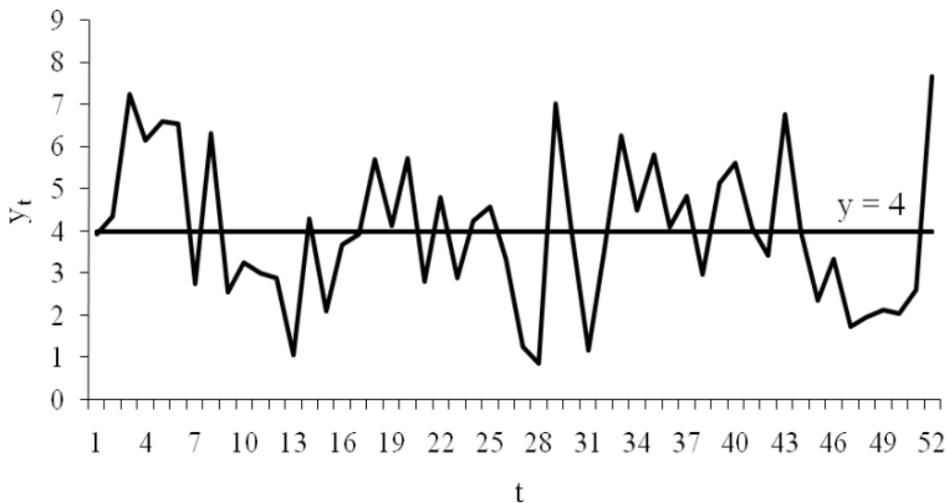


Рис. 1. График стационарного временного ряда

Экономические процессы во времени обычно не являются стационарными, так как содержат основную тенденцию развития, и описываются динамическими временными рядами. На рис. 2 график временного ряда располагается относительно наклонной прямой (тренда), характеризующей тенденцию.

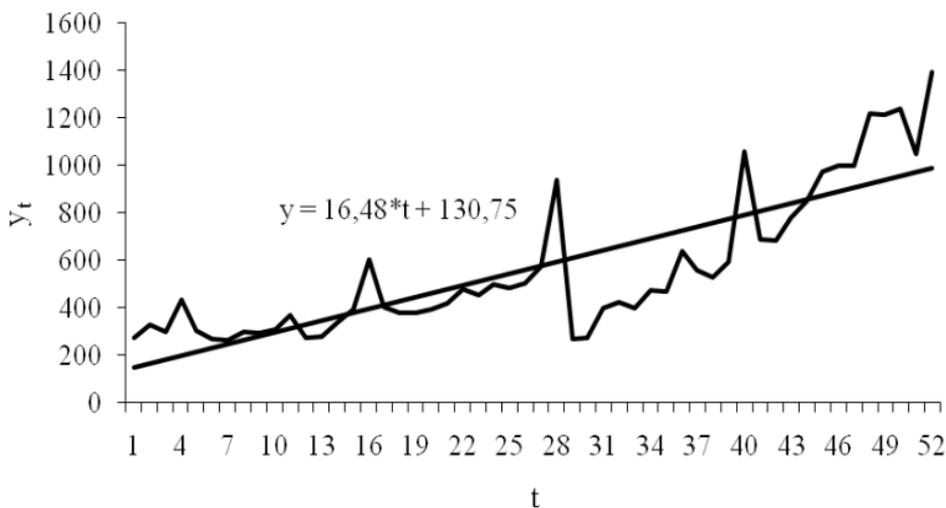


Рис. 2. График динамического временного ряда

5. По числу показателей выделяют *изолированные (одномерные) и компонентные (многомерные) временные ряды*. Если проводится анализ во времени одного показателя, то временной ряд изолированный. Например, объем произведенных предприятием товаров (помесячно) представляет собой одномерный временной ряд (табл. 1.1).

Таблица 1.1. Одномерный временной ряд

Момент времени (дата)	Январь 2004 г.	Февраль 2004 г.	Март 2004 г.	Апрель 2004 г.
Номер момента времени t	1	2	3	4	...	$n - 1$	n
Значение показателя y_t	y_1	y_2	y_3	y_4	...	y_{n-1}	y_n

В многомерном ряду представлена динамика нескольких показателей, характеризующих одно экономическое явление (процесс). Например, характеристики деятельности предприятия, наблюдаемые ежемесячно, можно представить в виде многомерного ряда (табл. 1.2). В качестве показателей производственной деятельности можно взять:

Y – производительность труда;

X_1 – удельный вес покупных изделий;

X_2 – размер премий и вознаграждений на одного работника;

X_3 – среднегодовая численность ППП;

X_4 – среднегодовой фонд заработной платы ППП;

X_5 – объем непроизводственных расходов.

Таблица 1.2. Многомерный временной ряд

Момент времени (дата)	Номер момента времени t	Значения показателей				
		Y	X_1	X_2	...	X_5
Январь 2004 г.	1	y_1	x_{11}	x_{12}	...	x_{15}
Февраль 2004 г.	2	y_2	x_{21}	x_{22}	...	x_{25}
Март 2004 г.	3	y_3	x_{31}	x_{32}	...	x_{35}
...
...	n	y_n	x_{n1}	x_{n2}	...	x_{n5}

В данном пособии изложена методика изучения абсолютных моментных одномерных временных рядов с равноотстоящими уровнями (далее – временные ряды).

1.3. Модели временных рядов

Значения уровней ряда формируются под влиянием различных факторов: тренда, сезонных и циклических закономерностей изучаемого процесса, а также случайных воздействий.

Тренд представляет собой *долговременную* систематическую линейную или нелинейную компоненту, закономерно изменяющуюся во времени.

Сезонная составляющая - это периодически повторяющаяся компонента, обусловленная временем года.

Циклическая составляющая - это периодически повторяющаяся компонента, обусловленная

технологическим процессом или экономическими закономерностями.

Эти три вида регулярных компонент часто имеются в рядах одновременно. Например, продажи компании могут возрасти из года в год (тренд), но при этом они могут содержать и сезонную составляющую (например, 30% годовых продаж приходится на январь и только 5% - на июль). На объем продаж также может влиять покупательная способность населения (циклическая составляющая), определяемая закономерностями в социально-экономической сфере.

Необязательно, чтобы в формировании значений уровней ряда участвовали все четыре вида факторов (табл. 1.3). Однако случайная составляющая присутствует всегда.

Таблица 1.3. Факторы, влияющие на значение уровней временного ряда

Компонента	Классификация	Определение
Тренд \hat{y}_t	Систематическая	Общая устойчивая долговременная тенденция
Сезонная s_t	Систематическая	Достаточно регулярные периодические изменения, происходящие в каждом 12-месячном периоде из года в год
Циклическая c_t	Систематическая	Повторяющиеся подъемы и спады, проходящие 4 фазы: пик, рецессия, депрессия, подъем
Случайная ε_t	Нерегулярная	Непредвиденные воздействия внешних факторов

Модель временного ряда, которая представляет каждое значение уровня ряда как сумму указанных выше компонент, называется *аддитивной*. Согласно этой модели любое значение временного ряда представляется в виде:

$$y_t = \hat{y}_t + s_t + c_t + \varepsilon_t, \quad (1.1)$$

где y_t - значение уровня временного ряда, а \hat{y}_t , s_t , c_t , ε_t - соответственно значения трендовой, сезонной, циклической и случайной компонент в любой точке ряда.

Аддитивная модель применяется в тех случаях, когда анализируемый временной ряд имеет приблизительно одинаковую амплитуду колебаний на протяжении всей длины ряда (рис. 3).

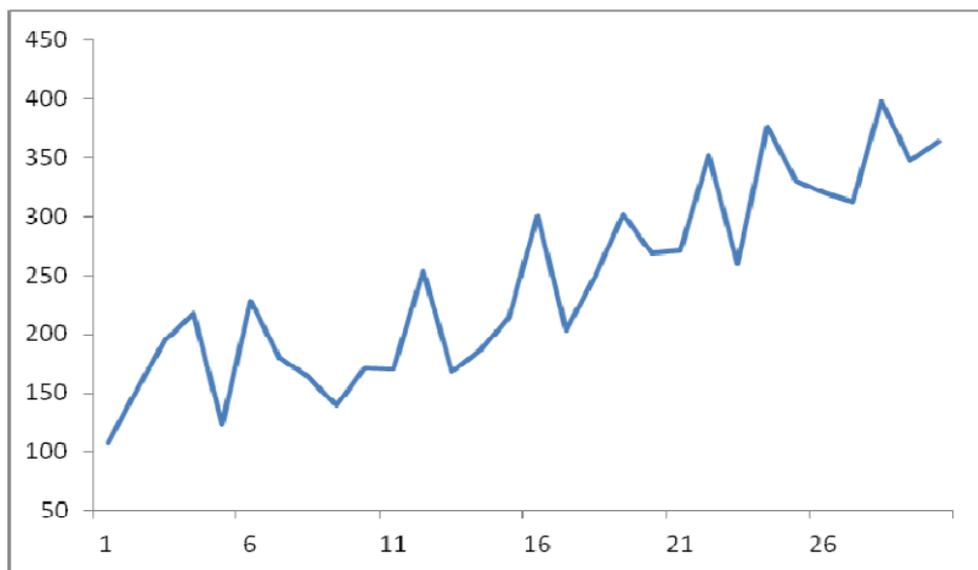


Рис. 3. Постоянная амплитуда колебаний уровней ряда

Наиболее фундаментальной является мультипликативная модель временного ряда, широко используемая при анализе ежемесячных, ежеквартальных и ежегодных данных и потому чаще всего применяемая в экономических исследованиях.

В *мультипликативной* модели временного ряда значения его уровней представляется в виде произведения:

$$y_t = \hat{y}_t \times s_t \times c_t \times \varepsilon_t . \quad (1.2)$$

Мультипликативная модель применяется в тех случаях, когда анализируемый временной ряд имеет возрастающую (уменьшающуюся) амплитуду колебаний на протяжении всей длины ряда (рис. 4).

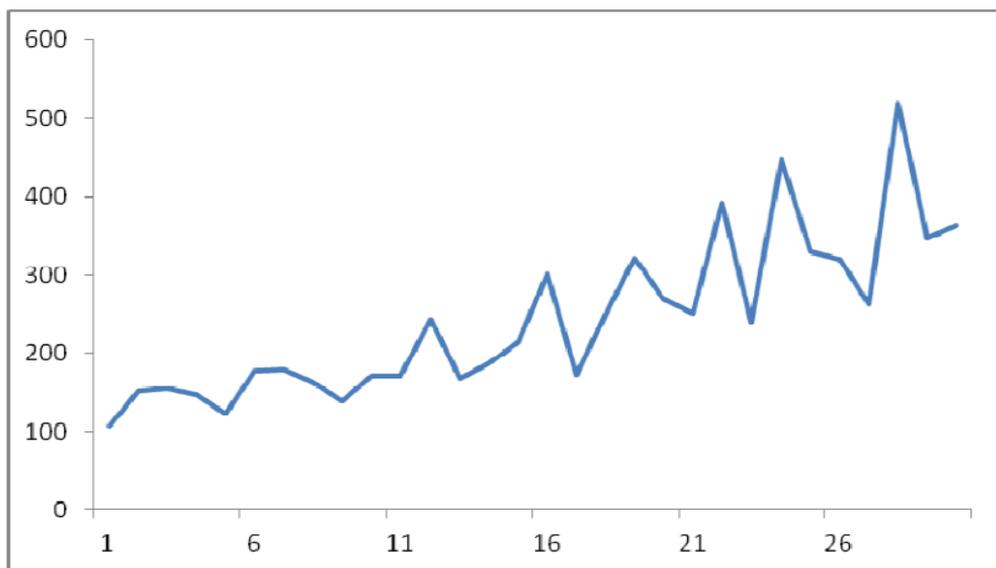


Рис. 4. Непостоянная амплитуда колебаний уровней ряда

2. Общий анализ временного ряда

2.1. Графический анализ временного ряда

Анализ временного ряда начинается с построения и анализа его графика.

Пример

Имеются месячные значения показателя «Среднедушевые денежные доходы» (руб.) для некоторого региона России за 1996–2001 гг. (табл. 2.1) (далее – «Денежные доходы»). Провести графический анализ ряда.

Таблица 2.1. Исходные данные

Период			Денежные доходы, руб., (y _t)	Период			Денежные доходы, руб., (y _t)
Год	Месяц.	t		Год	Месяц.	t	
1996	январь	1	407,1	1999	январь	37	727,3
1996	февраль	2	451,9	1999	февраль	38	868,7
1996	март	3	495,3	1999	март	39	965,2
1996	апрель	4	516,9	1999	апрель	40	1041,2
1996	май	5	523,8	1999	май	41	954,3
1996	июнь	6	567,7	1999	июнь	42	1058,5
1996	июль	7	479,7	1999	июль	43	1150,0
1996	август	8	463,7	1999	август	44	1072,0
1996	сентябрь	9	439,7	1999	сентябрь	45	1053,0
1996	октябрь	10	441,2	1999	октябрь	46	1539,9
1996	ноябрь	11	470,3	1999	ноябрь	47	1507,3
1996	декабрь	12	543,7	1999	декабрь	48	2013
1997	январь	13	468,1	2000	январь	49	1345,3
1997	февраль	14	486,6	2000	февраль	50	1587,0
1997	март	15	514,4	2000	март	51	1655,2
1997	апрель	16	601,0	2000	апрель	52	1723,0
1997	май	17	503,4	2000	май	53	1701,5

Период			Денежные доходы, руб., (y _t)	Период			Денежные доходы, руб., (y _t)
Год	Месяц.	t		Год	Месяц.	t	
1997	июнь	18	548,9	2000	июнь	54	1923,9
1997	июль	19	601,5	2000	июль	55	1852,6
1997	август	20	568,6	2000	август	56	1987,6
1997	сентябрь	21	571,3	2000	сентябрь	57	2003,3
1997	октябрь	22	632,3	2000	октябрь	58	1971,8
1997	ноябрь	23	589,7	2000	ноябрь	59	2027,0
1997	декабрь	24	796,3	2000	декабрь	60	2492,7
1998	январь	25	530,2	2001	январь	61	1793,8
1998	февраль	26	620,0	2001	февраль	62	2146,9
1998	март	27	612,3	2001	март	63	2422,3
1998	апрель	28	697,8	2001	апрель	64	2380,8
1998	май	29	647,9	2001	май	65	2315,4
1998	июнь	30	663,2	2001	июнь	66	2632,8
1998	июль	31	716,0	2001	июль	67	2584,1
1998	август	32	736,0	2001	август	68	2606,7
1998	сентябрь	33	731,7	2001	сентябрь	69	2547,5
1998	октябрь	34	674,1	2001	октябрь	70	2485,6
1998	ноябрь	35	712,9	2001	ноябрь	71	2611,8
1998	декабрь	36	1059,9	2001	декабрь	72	3406,5

Для проведения расчетов исходный ряд запишем в виде табл. 2.2.

Таблица 2.2. Временной ряд

t	y _t	t	y _t	t	y _t
1	407,1	25	530,2	49	1345,3
2	451,9	26	620,0	50	1587,0
3	495,3	27	612,3	51	1655,2
4	516,9	28	697,8	52	1723,0
5	523,8	29	647,9	53	1701,5
6	567,7	30	663,2	54	1923,9
7	479,7	31	716,0	55	1852,6
8	463,7	32	736,0	56	1987,6
9	439,7	33	731,7	57	2003,3

t	y_t	t	y_t	t	y_t
10	441,2	34	674,1	58	1971,8
11	470,3	35	712,9	59	2027,0
12	543,7	36	1059,9	60	2492,7
13	468,1	37	727,3	61	1793,8
14	486,6	38	868,7	62	2146,9
15	514,4	39	965,2	63	2422,3
16	601,0	40	1041,2	64	2380,8
17	503,4	41	954,3	65	2315,4
18	548,9	42	1058,5	66	2632,8
19	601,5	43	1150,0	67	2584,1
20	568,6	44	1072,0	68	2606,7
21	571,3	45	1053,0	69	2547,5
22	632,3	46	1539,9	70	2485,6
23	589,7	47	1507,3	71	2611,8
24	796,3	48	2013,0	72	3406,5

График временного ряда изучаемого показателя «Денежные доходы» изображен на рис. 5.



Рис. 5. График временного ряда

На рисунке видны периодические колебания графика. Наиболее выраженные колебания повторяются через каждые 12 моментов времени, что в нашем случае соответствует 12 месяцам. Между этими 12 месяцами так же есть незначительные колебания, повторяющиеся с определенной периодичностью. Таким образом, можно сделать предположение, что данный временной ряд не случаен и эти колебания – годовые, полугодовые и квартальные изменения доходов. Также можно заметить, что, начиная с $t=32$, график возрастает быстрее, а амплитуда колебаний увеличивается.

2.2. Проверка гипотезы о случайности тренда

Методы и методики анализа стационарных и нестационарных рядов существенно различаются. Визуальный анализ графика ряда не всегда позволяет с уверенностью сделать вывод о наличии (отсутствии) тенденции в изучаемых данных. Поэтому исследование временного ряда начинается с решения вопроса о его стационарности.

Одним из возможных вариантов для этого служит проверка гипотезы H_0 о случайности (отсутствии) тренда временного ряда при конкурирующей гипотезе H_1 о неслучайности (наличии) тренда, основанная на сравнении средних значений первой и второй половины ряда по статистике (расчетному значению критерия)

$$t_{расч} = \frac{\overline{y_n} - \overline{y_b}}{\sqrt{(n_n - 1)S_n^2 + (n_b - 1)S_b^2}} \cdot \sqrt{\frac{n_n \cdot n_b \cdot (n_n + n_b - 2)}{n_n + n_b}}, \quad (2.1)$$

где $\overline{y_n}$, $\overline{y_b}$ – средние значения первой и второй половины ряда, имеющих длины n_n , n_b ($n_n + n_b = n$); s_n^2 , s_b^2 – дисперсии первой и второй половины ряда.

Средние значения и дисперсии определяются по формулам

$$\overline{y_n} = \frac{1}{n_n} \sum_{t=1}^{n_n} y_t \quad \text{и} \quad \overline{y_b} = \frac{1}{n_b} \sum_{t=n_n+1}^n y_t, \quad (2.2)$$

$$s_n^2 = \frac{1}{n_n} \sum_{t=1}^{n_n} (y_t - \overline{y_n})^2 \quad \text{и} \quad s_b^2 = \frac{1}{n_b} \sum_{t=n_n+1}^n (y_t - \overline{y_b})^2. \quad (2.3)$$

Значение $|t_{\text{расч}}|$ сравнивается с критическим значением распределения Стьюдента $t_{\text{кр}}(\alpha, df)$ с $df = n_n + n_b - 2$ степенями свободы и уровнем значимости α .

В случае, если $|t_{\text{расч}}| < t_{\text{кр}}$ гипотеза H_0 о случайности тренда временного ряда принимается и ряд считается стационарным.

В противном случае ($|t_{\text{расч}}| > t_{\text{кр}}$) – гипотеза H_0 отвергается, что свидетельствует о значимости различия средних первой и второй половины ряда и неслучайности (наличии) тренда. Другими словами, ряд имеет тенденцию.

Для проверки гипотезы о наличии тренда также можно воспользоваться критерием серий, методом Фостера–Стюарта и др.

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1996 г. по декабрь 2001 г. проверить гипотезу о случайности тренда.

Для выяснения наличия динамики в изучаемом временном ряду проверим статистический критерий о равенстве двух средних.

Найдем статистику критерия по формуле (2.1):

$$t_{расч} = \frac{\bar{y}_n - \bar{y}_e}{\sqrt{(n_n - 1) \cdot s_n^2 + (n_e + 1) \cdot s_e^2}} \cdot \sqrt{\frac{n_n \cdot n_e \cdot (n_n + n_e - 2)}{n_n + n_e}},$$

где $n = 72$ – длина всего ряда, $n_{II} = 36$ – длина первой половины ряда, $n_{B} = 36$ – длина второй половины ряда, $n_{II} + n_{B} = 36 + 36 = 72$;

\bar{y}_{II} – средняя первой половины ряда,

$$\bar{y}_{II} = \frac{1}{n_{II}} \cdot \sum_{t=1}^{n_{II}} y_t = 585,70;$$

\bar{y}_{B} – средняя второй половины ряда,

$$\bar{y}_{B} = \frac{1}{n_{B}} \cdot \sum_{t=n+1}^n y_t = 1837,65;$$

s_{II}^2 – дисперсия первой половины ряда,

$$s_{\Pi}^2 = \frac{1}{n_{\Pi}} \cdot \sum_{t=1}^{n_{\Pi}} (y_t - \bar{y}_{\Pi})^2 = 16336,98;$$

s_B^2 – дисперсия второй половины ряда,

$$s_B^2 = \frac{1}{n_B} \cdot \sum_{t=n_{\Pi}+1}^n (y_t - \bar{y}_B)^2 = 414185,50.$$

$$t_{расч} = \frac{585,70 - 1837,65}{\sqrt{(36-1) \cdot 16336,98 + (36-1) \cdot 414185,50}} \times \sqrt{\frac{36 \cdot 36 \cdot (36 + 36 - 2)}{36 + 36}} = -11,45$$

Значение $|t_{расч}| = |-11,45| = 11,45$ сравним с критическим значением распределения Стьюдента, зависящим от уровня значимости $\alpha = 0,05$ и числа степеней свободы $df = n_{\Pi} + n_B - 2 = 70$.

Табличное (критическое) значение $t_{крит}$, найдем по таблице критических значений распределения Стьюдента (табл. Приложения 2, двухсторонняя критическая область) по уровню значимости α и числу степеней свободы df , или с помощью встроенной функции Excel «СТЮДРАСПОБР». $t_{крит}(0,05;70) = 1,99$.

Здесь и далее критические значения критериев найдены с помощью встроенных функций Excel.

В нашем случае $|t_{расч}| > t_{крит}$, что свидетельствует о значимости различия средних первой и второй половины ряда. Наличие тренда неслучайно. Ряд имеет тенденцию, т.е. не является стационарным.

2.3 Численное сглаживание уровней временного ряда. Метод скользящей средней

Сглаживание временных рядов проводится по следующим 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.

Сглаживание позволяет устранить случайные колебания уровней.

Некоторые методы анализа и прогнозирования требуют в качестве предварительного условия сглаживание временного ряда.

Методы сглаживания применяются для прогнозирования экономических показателей.

Существующие методы сглаживания делят на две группы:

1. *Аналитическое выравнивание* - сглаживание с использованием кривой (функции), проведенной относительно фактических значений ряда так, чтобы эта

кривая отображала тенденцию, присущую ряду и одновременно освобождала его от мелких незначительных колебаний. Такие кривые называют еще трендами (кривыми роста), и они используются главным образом для прогнозирования экономических показателей.

2. Методы *статистического сглаживания*. При использовании этих методов производится сглаживание каждого отдельного уровня ряда с использованием фактических значений соседних с ним уровней.

Рассмотрим сглаживание временных рядов методом скользящей средней.

Метод *простого* скользящего среднего используется обычно в тех случаях, когда график временного ряда близок к прямой линии, поскольку при этом динамика исследуемого явления не искажается ($p = 1$).

1. Определяется интервал сглаживания, т.е. число входящих в него уровней m ($m < n$). Число m определяют по следующим правилам:

Если нужно сгладить мелкие беспорядочные колебания, то значение m по возможности увеличивают, если же нужно сгладить циклически повторяющиеся колебания, то m уменьшают.

В качестве m рекомендуется брать нечетные числа.

2. Вычисляются средние значения уровней, входящих в интервал сглаживания.

$$\tilde{y}_t = \frac{\sum_{t=l}^{t+l} y_t}{m}, \quad (2.4)$$

где \tilde{y}_t – среднее значение уровней, входящих в интервал сглаживания, которое одновременно является сглаженным значением уровня, находящегося в середине интервала сглаживания. $l = (m - 1)/2$.

3. Производится сдвиг интервала сглаживания на одну точку вниз. Вычисляется среднее сглаженное значение для момента $t+1$, затем снова производится сдвиг вычисления и т.д.

Например, если сглаживание ряда осуществляется по $m = 5$ точкам, то для $p = 1$ (линейное сглаживание) новое сглаженное значение уровня ряда вычисляется по формуле

$$\tilde{y}_t = \frac{1}{5} (y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2}), \quad (2.5)$$

где y_t , \tilde{y}_t – заданное и новое сглаженное значения уровня ряда ($t = 3, 4, \dots, n - 2$).

4. В результате такой итерационной процедуры получаем $n - (m - 1)$ сглаженных уровней ряда. Не сглаживаются (теряются) l первых и l последних уровней исходного временного ряда, что является существенным недостатком данного метода.

При $p = 1$ для вычисления сглаженных первых и последних $(m - 1)/2$ значений ряда (при $m = 5$

вычисляются два первых и два последних члена) можно использовать следующие формулы:

$$\begin{aligned}\tilde{y}_1 &= \frac{1}{5}(3y_1 + 2y_2 + y_3 - y_5), \\ \tilde{y}_2 &= \frac{1}{10}(4y_1 + 3y_2 + 2y_3 + y_5), \\ \tilde{y}_{n-1} &= \frac{1}{10}(4y_n + 3y_{n-1} + 2y_{n-2} + y_{n-4}), \\ \tilde{y}_n &= \frac{1}{5}(3y_n + 2y_{n-1} + y_{n-2} - y_{n-4}).\end{aligned}\tag{2.6}$$

В том случае, когда тренд ряда имеет явно нелинейный характер (график ряда близок к кривой линии - полиному второго, третьего или более высокого порядка) и желательно сохранить незначительные колебания в динамике значений, этот метод не используется, так как его применение может привести к значительным искажениям исследуемого процесса. В таких случаях используется метод *взвешенной* скользящей средней.

Алгоритм метода *взвешенной* скользящей средней практически такой же, как и алгоритм метода простой скользящей средней. Различие состоит только в том, что каждому уровню, входящему в интервал сглаживания, передается определенный весовой коэффициент. Значение этого коэффициента зависит от порядка выбранного полинома l , от интервала сглаживания m и от местоположения уровня в интервале сглаживания. Степень полинома и число точек сглаживания выбираются из общих соображений, включая существо решаемой задачи, и подбора степени по пробным кратковременным прогнозам.

Для определения значений весовых коэффициентов используется метод наименьших квадратов, с помощью которого по m членам ряда строятся полиномы выбранной степени p , начиная с первого и т.д. членов ряда.

Если в качестве наилучшего выбран второй порядок полинома ($p = 2$, квадратичное сглаживание) и $m = 5$, то весовые коэффициенты будут следующие: $1/35$ (-3, 12, 17, 12, -3). Коэффициенты симметричны относительно центрального члена и сумма коэффициентов с учетом общего множителя равна единице.

Т.е. новое сглаженное значение уровня ряда находится по формуле

$$\tilde{y}_t = \frac{1}{35}(-3y_{t-2} + 12y_{t-1} + 17y_t + 12y_{t+1} - 3y_{t+2}). \quad (2.7)$$

Для вычисления сглаженных первых и последних $(m - 1)/2$ значений ряда (при $m = 5$ вычисляются два первых и два последних члена) используются следующие формулы:

$$\begin{aligned} \tilde{y}_1 &= \frac{1}{35}(31y_1 + 9y_2 - 3y_3 - 5y_4 + 3y_5), \\ \tilde{y}_2 &= \frac{1}{35}(9y_1 + 13y_2 + 12y_3 + 6y_4 - 5y_5), \\ \tilde{y}_{n-1} &= \frac{1}{35}(9y_n + 13y_{n-1} + 12y_{n-2} + 6y_{n-3} - 5y_{n-4}), \\ \tilde{y}_n &= \frac{1}{35}(31y_n + 9y_{n-1} - 3y_{n-2} - 5y_{n-3} + 3y_{n-4}). \end{aligned} \quad (2.8)$$

Недостаток этого метода такой же, как и у метода простой скользящей средней – не сглаживается (теряется) $(m - 1)/2$ первых и $(m - 1)/2$ последних уровней исходного временного ряда.

Кроме того, существуют методы, при использовании которого потерь уровней не происходит. К ним относится метод экспоненциального сглаживания.

Процедура численного сглаживания может применяться последовательно несколько раз.

После вычисления сглаженных значений ряда строится его графическое изображение (рис. 6).

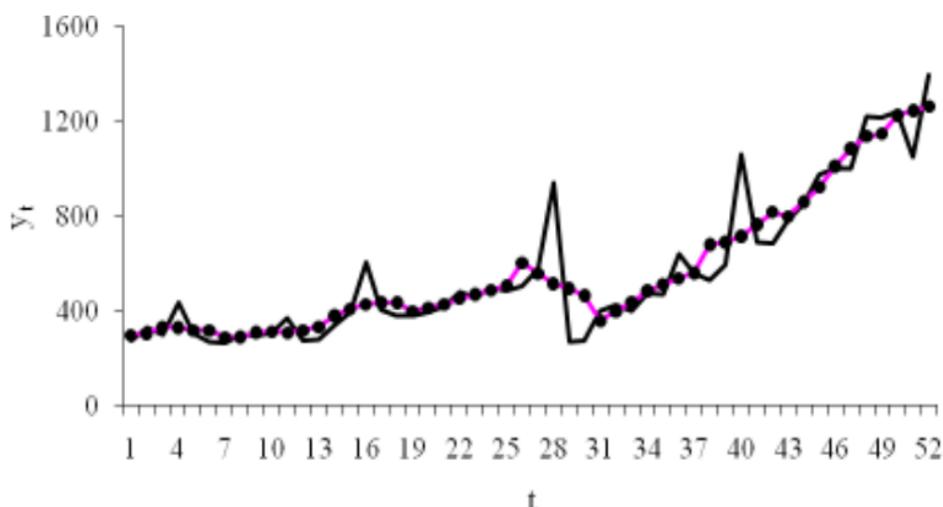


Рис. 6. Графики исходного и сглаженного ряда

Аналогичным способом производится сглаживание ряда по любому нечетному числу m членов ряда. В случае отсутствия необходимых формул, вычисление первых и последних членов сглаженного ряда не производится и в соответствующих строках расчетной таблицы ставится прочерк.

При сглаживании по четному числу членов ряда сначала вычисляются средние значения m уровней ряда, которые затем центрируются, т.е. в качестве сглаженных значений принимаются средние значения двух рядом стоящих средних.

В общем случае (линейное сглаживание, $p = 1$) расчет сглаженных значений \tilde{y}_t в случае нечетной длины интервала сглаживания $m = 2g + 1$ осуществляется по формуле

$$\tilde{y}_t = \frac{y_{t-g} + \dots + y_{t-1} + y_t + y_{t+1} + \dots + y_{t+g}}{2g + 1}, \quad (2.9)$$

$t = g+1, \dots, n-g$

(для $m = 3$ значение $g = 1$, для $m = 5$ значение $g = 2$).

Расчет сглаженных значений \tilde{y}_t в случае четной длины интервала сглаживания $m = 2g$ производится по формуле

$$\tilde{y}_t = \frac{\frac{1}{2}y_{t-g} + \dots + y_{t-1} + y_t + y_{t+1} + \dots + \frac{1}{2}y_{t+g}}{2g + 1}, \quad (2.10)$$

$t = g+1, \dots, n-g$

(для $m = 4$ значение $g = 2$).

Расчет сглаженных значений для незаполненных уровней ряда (уровни в начале и конце ряда), в котором отсутствует цикличность, можно произвести на основании

средних приростов. Сглаженные значения в начале временного ряда рассчитываются путем последовательного вычитания среднего прироста

$$\Delta y_{нач} = \frac{y_{2g+1} - y_1}{2g}$$

на первом активном участке из первого доступного сглаженного значения. Сглаженные значения в конце временного ряда рассчитываются путем последовательного прибавления среднего прироста

$$\Delta y_{кон} = \frac{y_n - y_{n-2g}}{2g}$$

на последнем активном участке к последнему доступному сглаженному значению. Под первым и последним активными интервалами понимаются интервалы, значения которых используются при расчетах первого и последнего усредненного значения.

В случае $m = 3$ ($g = 1$) первый активный участок включает в себя три первых уровней ряда, средний прирост

$$\Delta y_{нач} = \frac{y_3 - y_1}{2},$$

последний активный участок – три последних уровня ряда,

$$\Delta y_{кон} = \frac{y_n - y_{n-2}}{2}.$$

В случае $m = 5$ ($g = 2$) первый активный участок включает в себя пять первых уровней ряда, средний прирост

$$\Delta y_{нач} = \frac{y_5 - y_1}{2 \cdot 2},$$

а последний активный участок – пять последних уровней ряда и

$$\Delta y_{кон} = \frac{y_n - y_{n-4}}{2 \cdot 2}.$$

Поскольку расчет сглаженных значений для $m = 5$ и $m = 4$ производится по значениям одних и тех же интервалов, то средние приросты для этих значений параметра m также одинаковы.

В случае $m = 3$ необходимо рассчитать сглаженные значения только для первого и последнего уровней ряда:

$$\tilde{y}_1 = \tilde{y}_2 - \Delta y_{нач}; \quad \tilde{y}_n = \tilde{y}_{n-1} + \Delta y_{кон}. \quad (2.11)$$

В случае $m = 5$ необходимо рассчитать сглаженные значения для двух первых и двух последних уровней ряда:

$$\tilde{y}_2 = \tilde{y}_3 - \Delta y_{нач}; \quad \tilde{y}_1 = \tilde{y}_2 - \Delta y_{нач}; \quad (2.12)$$

$$\tilde{y}_{n-1} = \tilde{y}_{n-2} + \Delta y_{кон}; \quad \tilde{y}_n = \tilde{y}_{n-1} + \Delta y_{кон}. \quad (2.13)$$

Аналогично в случае $m = 4$ необходимо рассчитать сглаженные значения для двух первых и двух последних уровней ряда по формулам (2.12; 2.13).

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1996 г. по декабрь 2001 г. провести линейное сглаживание методом скользящей средней по пяти точкам. На одном рисунке построить график временного ряда и график сглаженного ряда.

Проведем линейное сглаживание ряда по $m = 5$ точкам. Новое сглаженное значение уровня ряда \tilde{y}_t , для $t = 3, 4, \dots, n - 2$, вычисляется по формуле (2.5):

$$\tilde{y}_t = \frac{1}{5}(y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2}).$$

Для $t = 3$

$$\tilde{y}_3 = \frac{1}{5}(y_1 + y_2 + y_3 + y_4 + y_5) = 479 ;$$

для $t = 4$

$$\tilde{y}_4 = \frac{1}{5}(y_2 + y_3 + y_4 + y_5 + y_6) = 511,12$$

и т.д.

Для вычисления двух первых и двух последних сглаженных уровней ряда используем формулы:

$$\tilde{y}_1 = \frac{1}{5}(3y_1 + 2y_2 + y_3 - y_5) = 419,32 ,$$

$$\tilde{y}_2 = \frac{1}{10}(4y_1 + 3y_2 + 2y_3 + y_5) = 449,85 ,$$

$$\tilde{y}_{n-1} = \frac{1}{10}(4y_n + 3y_{n-1} + 2y_{n-2} + y_{n-4}),$$

$$\tilde{y}_{71} = \frac{1}{10}(4y_{72} + 3y_{71} + 2y_{70} + y_{68}) = 2903,93 ;$$

$$\tilde{y}_n = \frac{1}{5}(3y_n + 2y_{n-1} + y_{n-2} - y_{n-4});$$

$$\tilde{y}_{72} = \frac{1}{5}(3y_{72} + 2y_{71} + y_{70} - y_{68}) = 3064,4.$$

Исходный и сглаженный ряд представлены в табл. 2.3 и на рис. 7.

Таблица 2.3. Расчет сглаженных уровней ряда

t	y_t	\tilde{y}_t	t	y_t	\tilde{y}_t
1	407,1	419,32	37	727,3	866,8
2	451,9	449,85	38	868,7	932,46
3	495,3	479,00	39	965,2	911,34
4	516,9	511,12	40	1041,2	977,58
5	523,8	516,68	41	954,3	1033,84
6	567,7	510,36	42	1058,5	1055,20
7	479,7	494,92	43	1150,0	1057,56
8	463,7	478,40	44	1072,0	1174,68
9	439,7	458,92	45	1053,0	1264,44
10	441,2	471,72	46	1539,9	1437,04
11	470,3	472,60	47	1507,3	1491,70
12	543,7	481,98	48	2013,0	1598,50
13	468,1	496,62	49	1345,3	1621,56
14	486,6	522,76	50	1587,0	1664,70
15	514,4	514,70	51	1655,2	1602,40
16	601,0	530,86	52	1723,0	1718,12
17	503,4	553,84	53	1701,5	1771,24
18	548,9	564,68	54	1923,9	1837,72
19	601,5	558,74	55	1852,6	1893,78

t	y_t	\tilde{y}_t	t	y_t	\tilde{y}_t
20	568,6	584,52	56	1987,6	1947,84
21	571,3	592,68	57	2003,3	1968,46
22	632,3	631,64	58	1971,8	2096,48
23	589,7	623,96	59	2027,0	2057,72
24	796,3	633,70	60	2492,7	2086,44
25	530,2	629,70	61	1793,8	2176,54
26	620,0	651,32	62	2146,9	2247,30
27	612,3	621,64	63	2422,3	2211,84
28	697,8	648,24	64	2380,8	2379,64
29	647,9	667,44	65	2315,4	2467,08
30	663,2	692,18	66	2632,8	2503,96
31	716,0	698,96	67	2584,1	2537,30
32	736,0	704,20	68	2606,7	2571,34
33	731,7	714,14	69	2547,5	2567,14
34	674,1	782,92	70	2485,6	2731,62
35	712,9	781,18	71	2611,8	2903,93
36	1059,9	808,58	72	3406,5	3064,40

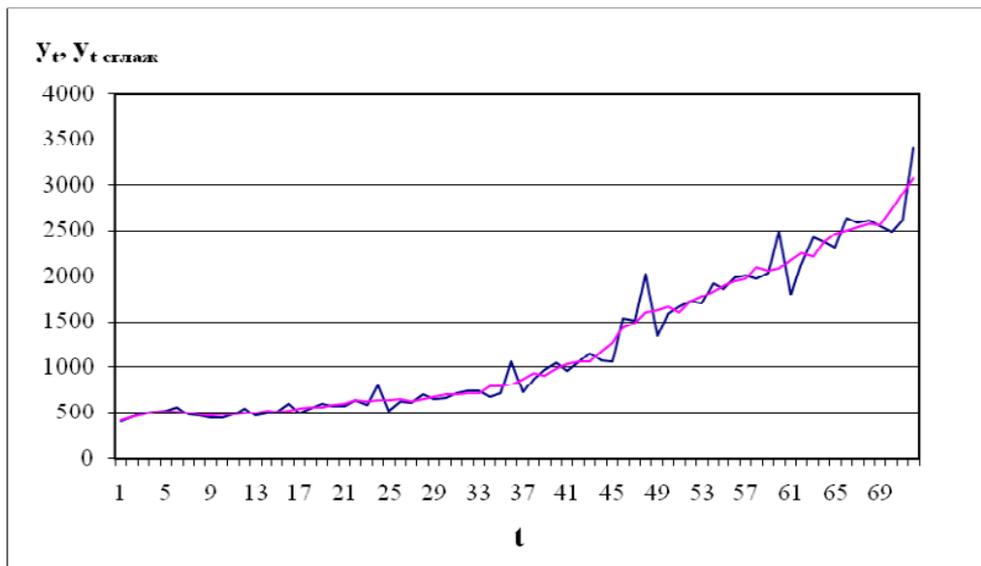


Рис. 7. Графики исходного и сглаженного ряда

2.4. Исследование структуры ряда.

Автокорреляция

При наличии во временном ряду тенденции и циклических колебаний значение каждого последующего уровня ряда зависят от значений предыдущих уровней. В этом случае говорят, что ряд имеет автокорреляцию.

Автокорреляционный анализ временного ряда позволяет установить степень зависимости последующих членов ряда от предыдущих и временной интервал, в течение которого эта зависимость статистически значима.

Автокорреляцией называется корреляция уровней временного ряда друг с другом, со сдвигом во времени на l тактов (лагом l). Количественно ее можно измерить с помощью парного коэффициента корреляции между уровнями исходного временного ряда и уровнями этого ряда, сдвинутыми на l шагов во времени.

Коэффициент автокорреляции первого порядка (лаг $l = 1$) находится по формуле

$$r_1 = \frac{\sum_{t=2}^n (y_t - \bar{y}_t)(y_{t-1} - \bar{y}_{t-1})}{\sqrt{\sum_{t=2}^n (y_t - \bar{y}_t)^2 \cdot \sum_{t=2}^n (y_{t-1} - \bar{y}_{t-1})^2}}, \quad (2.14)$$

где

$$\bar{y}_t = \frac{1}{n-1} \sum_{t=2}^n y_t; \quad \bar{y}_{t-1} = \frac{1}{n-1} \sum_{t=2}^n y_{t-1}.$$

Коэффициент автокорреляции l -го порядка находится по формуле

$$r_l = \frac{\sum_{t=l+1}^n (y_t - \bar{y}_t)(y_{t-l} - \bar{y}_{t-l})}{\sqrt{\sum_{t=l+1}^n (y_t - \bar{y}_t)^2 \cdot \sum_{t=l+1}^n (y_{t-l} - \bar{y}_{t-l})^2}}, \quad (2.15)$$

где

$$\bar{y}_t = \frac{1}{n-l} \sum_{t=l+1}^n y_t; \quad \bar{y}_{t-l} = \frac{1}{n-l} \sum_{t=l+1}^n y_{t-l}.$$

После вычисления коэффициентов автокорреляции необходимо проверить их статистическую значимость сравнением с критическими значениями коэффициента корреляции $r'_{\text{крит}}(\alpha, k)$. Критические значения берутся из таблицы критических значений корреляции по уровню значимости α и числу степеней свободы $k = n - l - 2$. Если $|r_l| < r'_{\text{крит}}$ то коэффициент корреляции r_l статистически незначим и выводы, сделанные по его значению, имеют вероятность ошибки, равную $1 - \alpha$.

Последовательность коэффициентов автокорреляции называют **автокорреляционной функцией** временного ряда. График зависимости ее значений от величины лага (порядка коэффициента автокорреляции) называется **коррелограммой** (корреограммой). Поскольку знаки коэффициентов автокорреляции при анализе не учитываются, коррелограмма обычно строится по их абсолютным значениям (рис. 8).

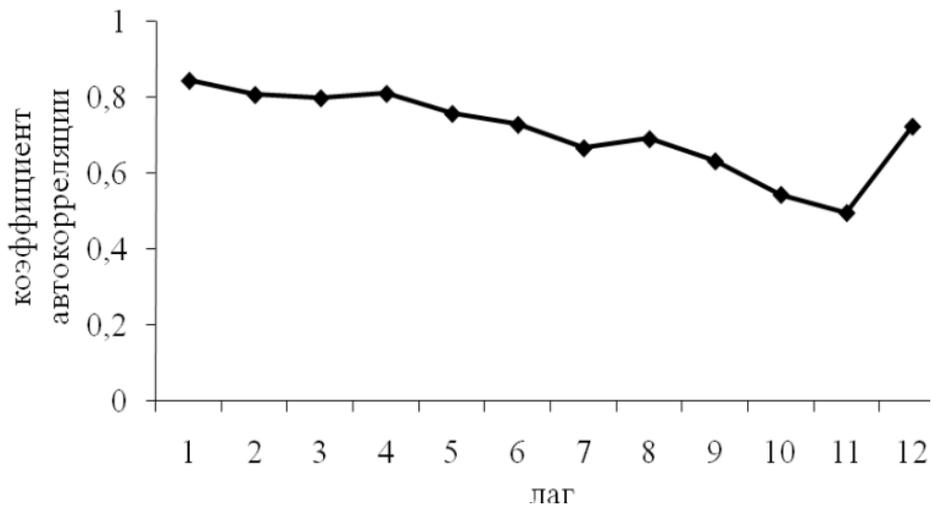


Рис. 8. Коррелограмма

Анализ коэффициентов автокорреляции производится на основании шкалы Чеддока (табл. 2.4).

Таблица 2.4. Шкала Чеддока

Значение $ r_l $	Характер тесноты связи между уровнями y_t и y_{t+l}
0 - 0,2	Связь практически отсутствует
0,2 - 0,5	Связь слабая
0,5 - 0,7	Связь средняя (умеренная)
0,7 - 0,95	Связь сильная (тесная)
0,95 - 1	Практически функциональная зависимость

Анализ автокорреляционной функции и коррелограммы позволяет определить лаг, при котором связь между текущими и предыдущими уровнями ряда наиболее тесная (сильная).

Если абсолютное значение коэффициента автокорреляции первого порядка $|r_1| > 0,7$, ряд содержит линейную тенденцию, если $|r_1| < 0,5$ - ряд содержит нелинейную тенденцию. В случае $0,5 < |r_1| < 0,7$ выбор вида уравнения тренда требует дополнительных исследований.

В случае, когда наибольшее абсолютное значение имеет коэффициент автокорреляции порядка $l = \tau > 1$ и при этом, $|r_\tau| > 0,7$, ряд содержит циклические колебания с периодом в τ моментов времени.

Возникают ситуации, когда $|r_l| > 0,7$ и $l \geq 2$, но сущность изучаемого процесса, а также вид графика ряда не позволяют сделать вывод о наличии цикличности. В таких случаях динамика описывается авторегрессионным уравнением либо уравнением с распределенным лагом.

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. вычислить коэффициенты автокорреляции. Проверить статистическую значимость коэффициентов автокорреляции. Построить коррелограмму.

Для расчета коэффициентов автокорреляции 1-го, 2-го, ..., 12-го порядка составим вспомогательную таблицу (табл. 2.5).

Таблица 2.5. Представление данных для расчета коэффициентов автокорреляции

t	y_t	y_{t-1}	y_{t-2}	y_{t-3}	y_{t-4}	y_{t-5}	y_{t-6}	y_{t-7}	y_{t-8}	y_{t-9}	y_{t-10}	y_{t-11}	y_{t-12}
1	727												
2	869	727											
3	965	869	727										
4	1041	965	869	727									
5	954	1041	965	869	727								
6	1059	954	1041	965	869	727							
7	1150	1059	954	1041	965	869	727						
8	1072	1150	1059	954	1041	965	869	727					
9	1053	1072	1150	1059	954	1041	965	869	727				
10	1540	1053	1072	1150	1059	954	1041	965	869	727			
11	1507	1540	1053	1072	1150	1059	954	1041	965	869	727		
12	2013	1507	1540	1053	1072	1150	1059	954	1041	965	869	727	
13	1345	2013	1507	1540	1053	1072	1150	1059	954	1041	965	869	727
14	1587	1345	2013	1507	1540	1053	1072	1150	1059	954	1041	965	869
15	1655	1587	1345	2013	1507	1540	1053	1072	1150	1059	954	1041	965
16	1723	1655	1587	1345	2013	1507	1540	1053	1072	1150	1059	954	1041
17	1702	1723	1655	1587	1345	2013	1507	1540	1053	1072	1150	1059	954
18	1924	1702	1723	1655	1587	1345	2013	1507	1540	1053	1072	1150	1059
19	1853	1924	1702	1723	1655	1587	1345	2013	1507	1540	1053	1072	1150
20	1988	1853	1924	1702	1723	1655	1587	1345	2013	1507	1540	1053	1072
21	2003	1988	1853	1924	1702	1723	1655	1587	1345	2013	1507	1540	1053
22	1972	2003	1988	1853	1924	1702	1723	1655	1587	1345	2013	1507	1540
23	2027	1972	2003	1988	1853	1924	1702	1723	1655	1587	1345	2013	1507
24	2493	2027	1972	2003	1988	1853	1924	1702	1723	1655	1587	1345	2013
25	1794	2493	2027	1972	2003	1988	1853	1924	1702	1723	1655	1587	1345
26	2147	1794	2493	2027	1972	2003	1988	1853	1924	1702	1723	1655	1587
27	2422	2147	1794	2493	2027	1972	2003	1988	1853	1924	1702	1723	1655
28	2381	2422	2147	1794	2493	2027	1972	2003	1988	1853	1924	1702	1723
29	2315	2381	2422	2147	1794	2493	2027	1972	2003	1988	1853	1924	1702
30	2633	2315	2381	2422	2147	1794	2493	2027	1972	2003	1988	1853	1924
31	2584	2633	2315	2381	2422	2147	1794	2493	2027	1972	2003	1988	1853
32	2607	2584	2633	2315	2381	2422	2147	1794	2493	2027	1972	2003	1988
33	2548	2607	2584	2633	2315	2381	2422	2147	1794	2493	2027	1972	2003
34	2486	2548	2607	2584	2633	2315	2381	2422	2147	1794	2493	2027	1972
35	2612	2486	2548	2607	2584	2633	2315	2381	2422	2147	1794	2493	2027
36	3407	2612	2486	2548	2607	2584	2633	2315	2381	2422	2147	1794	2493

Коэффициент автокорреляции l -го порядка найдем по формуле (2.15):

Фактически он является коэффициентом парной корреляции между соответствующими столбцами табл. 2.5 и находится с помощью встроенной функции Excel «КОРРЕЛ». Например, коэффициент автокорреляции второго порядка r_2 есть коэффициент парной корреляции между выделенными заливкой столбцами табл. 2.5.

На основе табл. 2.5 рассчитываются коэффициенты автокорреляции и заполняется табл. 2.6. Первые два столбца которой представляют собой табличное задание автокорреляционной функции. Ее график (коррелограмма), построенный по абсолютным значениям коэффициентов автокорреляции, изображен на рис. 9.

Таблица 2.6. Автокорреляционная функция.

Статистическая значимость

Лаг l	Модуль коэффициента автокорреляции r_l	Число степеней свободы k	Критическое значение корреляции $r'_{\text{крит}}$
1	2	3	4
1	0,90	33	0,344
2	0,89	32	0,349
3	0,89	31	0,355
4	0,90	30	0,361
5	0,87	29	0,367
6	0,90	28	0,374
7	0,84	27	0,381
8	0,88	26	0,388
9	0,86	25	0,396
10	0,80	24	0,404
11	0,72	23	0,413
12	0,95	22	0,423

Проверка статистической значимости найденных коэффициентов автокорреляции производится следующим образом.

Для каждого лага l вычисляется число степеней свободы $k = n - l - 2 = 36 - l - 2$ (указаны в 3-м столбце табл. 2.6).

Для уровня значимости $\alpha = 0,05$ находятся критические значения $r'_{\text{крит}}$ коэффициентов корреляции по таблице Приложения 3 (указаны в 4-м столбце табл. 2.6).

Если $|r_l| > r'_{\text{крит}}$ то коэффициент автокорреляции r_l статистически значим (если $|r_l| < r'_{\text{крит}}$ то коэффициент автокорреляции r_l статистически незначим).

Проверим выполнение этого неравенства графически. На рис. 9 добавим линию критических значений (прямая линия).

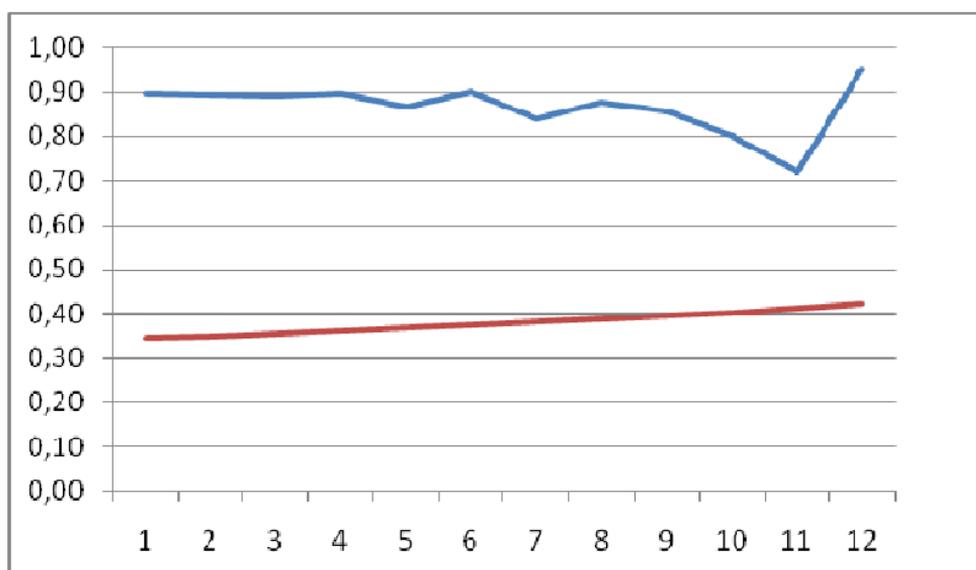


Рис. 9. Коррелограмма и линия критических значений

Для всех точек коррелограммы, лежащих выше линии критических значений, выполняется неравенство $|r_l| > r_{\text{крит}}^l$, следовательно соответствующие коэффициенты автокорреляции статистически значимы. В нашем случае все коэффициенты автокорреляции статистически значимы.

Поскольку $|r_1| = 0,90 \geq 0,70$, изучаемый ряд имеет линейную тенденцию, уравнение тренда – прямая $\hat{y}_t = b_0 + b_1 \cdot t$.

Как видно из табл. 2.6 и рис. 9, для $l = 12$ коэффициент автокорреляции r_{12} принимает максимальное значение, следовательно, данный ряд имеет цикличность с периодом в 12 моментов времени (т.е. годовой цикл на месячных данных). Длина цикла $\tau = 12$.

3. Моделирование временного ряда без учета сезонности

3.1. Виды уравнений тренда

В эконометрике построение функции (уравнения тренда) для моделирования тенденции временного ряда называют аналитическим выравнением временного ряда.

Основным содержанием метода аналитического выравнения временных рядов является расчет общей тенденции развития (тренда) как функции времени $\hat{y}_t = f(t)$, где \hat{y}_t - теоретические значения уровней ряда, вычисленные по соответствующему аналитическому выражению на момент времени t .

Простейшими функциями (уравнениями, формулами), выражающими тенденцию развития, являются следующие:

1. полиномиальная (p - степень полинома)

$$\hat{y}_t = b_0 + b_1 \cdot t + b_2 \cdot t^2 + \dots + b_p \cdot t^p;$$

2. линейная $\hat{y}_t = b_0 + b_1 \cdot t$ (частный случай

полиномиальной при $p = 1$), используется для описания процессов, развитие которых протекает во времени равномерно. Параметр b_0 интерпретируется как параметр начальных условий (значение уровня y_0 в нулевой момент времени $t = 0$), b_1 - как скорость роста значений уровней y_t ;

3. параболическая $\hat{y}_t = b_0 + b_1 \cdot t + b_2 \cdot t^2$ (частный случай полиномиальной при $p = 2$), используется для описания процессов, развитие которых характеризуется равноускоренным ростом (снижением) во времени. Параметр b_0 интерпретируется как параметр начальных условий (значение переменной y_0 в нулевой момент времени $t = 0$), b_1 - как скорость роста и b_2 - как ускорение роста значений уровней y_t ;

4. показательная $\hat{y}_t = b_0 \cdot b_1^t$;

5. экспоненциальная $\hat{y}_t = b_0 \cdot e^{b_1 t}$;

6. степенная $\hat{y}_t = b_0 \cdot t^{b_1}$;

7. гиперболическая $\hat{y}_t = b_0 + \frac{b_1}{t}$.

Параметры уравнений трендов (уравнение тренда представляет собой уравнение регрессии), как правило, определяются методом наименьших квадратов (в случае нелинейной функции – с предварительной линеаризацией). В качестве независимой переменной выступает время $t = 1, 2, \dots, n$, а в качестве зависимой переменной - фактические уровни временного ряда y_t .

На практике при выборе формы тренда обычно используют положения и выводы экономической теории, визуальный анализ графика ряда, а также результаты

исследования структуры ряда (автокорреляция уровней ряда, определение степени полиномиального тренда и др.).

На рис. 10 изображен линейный тренд, на рис. 19 - полиномиальный тренд.

В качестве критериев отбора наилучшей формы тренда часто используют наибольшее значение скорректированного коэффициента детерминации $R_{\text{скорр}}^2$ и наименьшее значение средней относительной ошибки аппроксимации \bar{A} .

3.2. Построение линейного тренда

Поведение многих монотонных временных рядов хорошо описывается линейной функцией, т.е. линейным трендом

$$\hat{y}_t = b_0 + b_1 \cdot t . \quad (3.1)$$

Параметры линейного тренда находятся методом наименьших квадратов по формулам:

$$b_1 = \frac{\sum_{t=1}^n t \cdot y_t - \frac{1}{n} \sum_{t=1}^n t \sum_{t=1}^n y_t}{\sum_{t=1}^n t^2 - \frac{1}{n} \left(\sum_{t=1}^n t \right)^2} , \quad (3.2)$$

$$b_0 = \bar{y}_t - b_1 \cdot \bar{t} . \quad (3.3)$$

Здесь \bar{y}_t , \bar{t} - средние уровней ряда y_t и времени t .

Для нахождения значений параметров b_0 , b_1 рекомендуется использовать пакет Excel «Анализ данных», инструмент «Регрессия».

На рис. 10 прямой линией изображен линейный тренд.

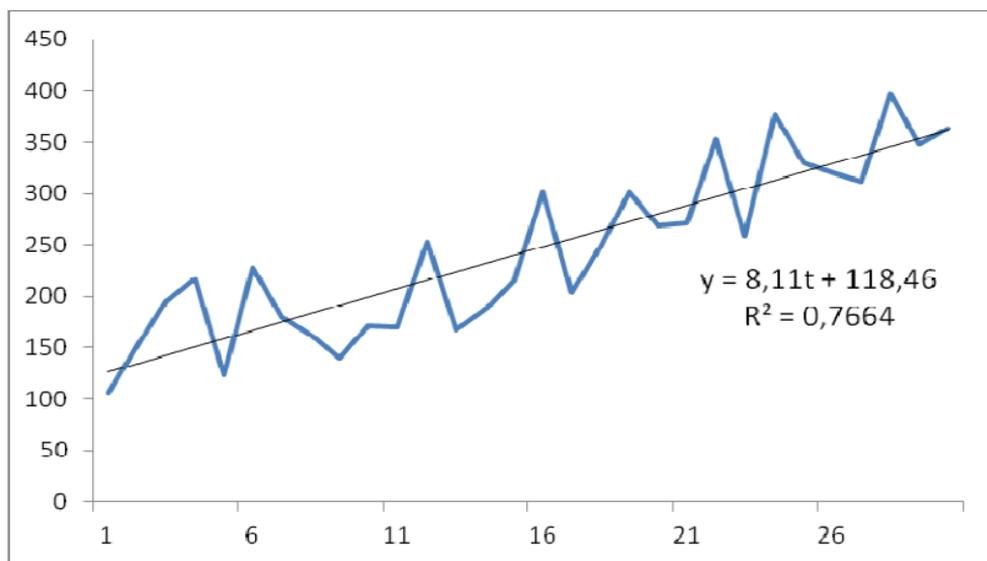


Рис. 10. Графики ряда и линейного тренда

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. построить линейный тренд.

Автокорреляционный анализ показал (стр. 41), что изучаемый ряд имеет линейную тенденцию, которая описывается уравнением линейного тренда $\hat{y}_t = b_0 + b_1 t$. Это уравнение является уравнением парной линейной регрессии, в котором в качестве независимой переменной выступает время t , а в качестве зависимой переменной — исходный ряд y_t .

Для нахождения параметров тренда b_0 и b_1 воспользуемся программой «Анализ данных», инструмент «Регрессия». Предварительно перенумеровав уровни ряда, начиная с 1 до 36. Результаты расчетов приведены в табл. 3.1.

Таблица 3.1. Результаты расчетов параметров и характеристик линейного тренда

ВЫВОД ИТОГОВ

<i>Регрессионная статистика</i>	
Множественный R	0,955
R-квадрат	0,911
Нормированный R-квадрат	0,908
Стандартная ошибка	194,708
Наблюдения	36

Дисперсионный анализ

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	1	13207515,35	13207515,35	348,38	1,93E-19
Остаток	34	1288977,12	37911,09		
Итого	35	14496492,47			

	<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	758,99	66,28	11,45	3,25E-13	624,29	893,68
t	58,31	3,12	18,66	1,93E-19	51,96	64,65

Уравнение тренда имеет вид

$$\hat{y}_t = 758,99 + 58,31 \cdot t \quad (t = 1, 2, \dots, 36).$$

Прямая линия тренда изображена на рис. 11.

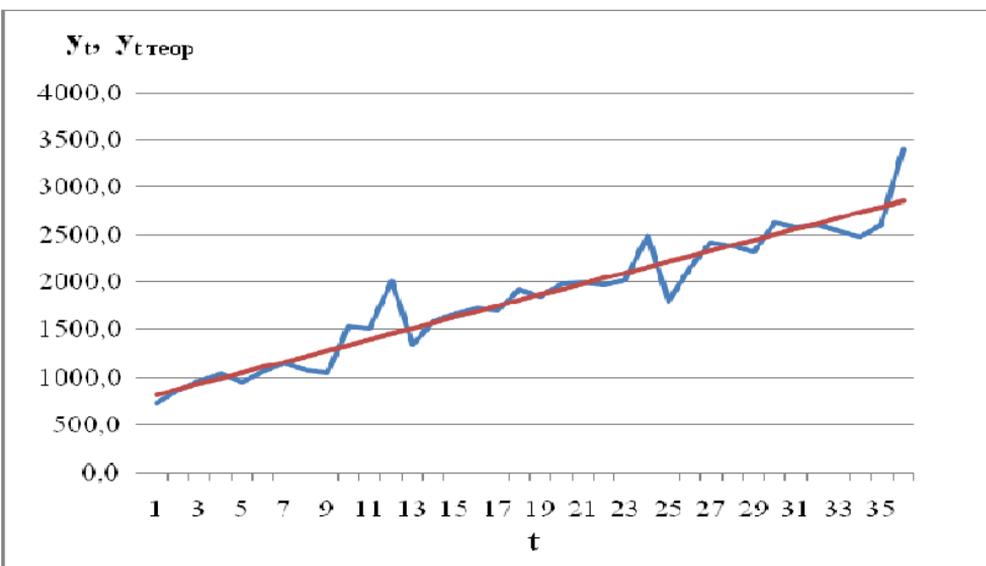


Рис. 11. График ряда и линия тренда

3.3. Оценка качества уравнения линейного тренда

Оценка качества уравнения линейного тренда производится аналогично оценке качества уравнения парной линейной регрессии.

1. Найденное уравнение линейного тренда графически изображается прямой линией (рис. 11), проходящей внутри ломаной линии – графика ряда.

2. Если $\bar{y}_t \approx \hat{y}_t$, арифметические расчеты выполнены верно.

3. Для оценки математической точности уравнения тренда можно воспользоваться средней относительной ошибкой аппроксимации

$$\bar{A} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \cdot 100\% \quad (3.4)$$

Принятие решения о точности уравнения основывается на табл. 3.2.

Таблица 3.2. Характеристика точности уравнения

Значение \bar{A} , %	Точность уравнения
менее 10	высокая
10 - 20	хорошая
20 - 50	удовлетворительная
более 50	неудовлетворительная

4. Проверка статистической значимости уравнения линейного тренда в целом осуществляется с помощью F-критерия Фишера.

Находится расчетное значение (статистика) критерия

$$F_{расч} = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{y}_t)^2}{\sum_{t=1}^n (\hat{y}_t - y_t)^2} \cdot \frac{n - m - 1}{m} \quad (3.5)$$

Здесь, n – длина ряда, $(m + 1)$ – число параметров уравнения тренда. В случае линейного тренда $m = 1$.

Табличное (критическое) значение $F_{табл}$, находится по таблице критических значений распределения Фишера-Снедекора (F-распределения) (табл. Приложения 1) по уровню значимости α и двум числам степеней свободы

$df_1 = m$ и $df_2 = n - m - 1$, или с помощью встроенной функции Excel «ФРАСПОБР».

Если $F_{расч} > F_{табл}$, то с вероятностью ошибки α уравнение линейного тренда в целом статистически значимо (адекватно).

В противном случае ($F_{расч} < F_{табл}$) уравнение в целом статистически незначимо.

5. Для проверки статистической значимости оценок параметров b_0, b_1 можно воспользоваться t-критерием Стьюдента.

Находятся расчетные значения критерия

$$t_{b_0} = \frac{b_0}{m_{b_0}}, \quad t_{b_1} = \frac{b_1}{m_{b_1}} \quad (3.6)$$

где средние квадратические ошибки параметров равны:

$$m_{b_0} = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{(n-2)} \cdot \frac{\sum_{i=1}^n t^2}{n \cdot \sum_{i=1}^n (t - \bar{t})^2}} \quad (3.7)$$

$$m_{b_1} = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{(n-2) \cdot \sum_{t=1}^n (t - \bar{t})^2}} \quad (3.8)$$

Теоретическое значение критерия $t_{\text{табл}}$ находится по таблице критических значений распределения Стьюдента (табл. Приложения 2) по уровню значимости α и числу степеней свободы $df = n - m - 1$, или с помощью встроенной функции Excel «СТЮДРАСПОБР».

Если $|t_{bj}| > t_{\text{табл}}$, то с вероятностью ошибки α оценка параметра уравнения регрессии b_j ($j = 0, 1$) статистически значима.

В противном случае ($|t_{bj}| < t_{\text{табл}}$) - b_j статистически незначима.

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. построен линейный тренд. Оценить качество уравнения тренда (найти среднюю относительную ошибку аппроксимации, проверить критерии Фишера и Стьюдента).

Оценка качества модели временного ряда (уравнение тренда найдено в предыдущем примере)

$$y_t = 758,99 + 58,31 \cdot t + \varepsilon_t$$

производится аналогично оценки качества линейной регрессионной модели.

1. Для проверки математической точности уравнения линейного тренда

$$\hat{y}_t = 758,99 + 58,31 \cdot t$$

найдем среднюю относительную ошибку аппроксимации по формуле (3.4) (расчеты см. в табл. 3.3).

$$\bar{A} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \cdot 100\% = \frac{1}{36} \cdot 2,566 \cdot 100\% = 7,13\% .$$

Так как $7,13\% < 10\%$ точность уравнения тренда высокая.

2. Проверку статистической значимости уравнения тренда в целом произведем с помощью F-критерия Фишера (3.5). Расчетные значения взяты из табл. 3.1.

$$F_{расч} = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \hat{y}_t)^2} \cdot \frac{n - m - 1}{m} = 348,38 ,$$

где $n = 36$ – длина ряда, $m = 1$ – число независимых переменных в уравнении тренда.

Табличное значение критерия Фишера с $df_1 = p = 1$ и $df_2 = n - m - 1 = 36 - 1 - 1 = 34$ степенями свободы при уровне значимости $\alpha = 0,05$

$$F_{табл}(0,05; 1; 34) = 4,13 .$$

Поскольку $F_{расч} > F_{табл}$, уравнение тренда в целом статистически значимо.

3. Проверку статистической значимости параметров уравнения тренда произведем с помощью критерия

Стьюдента (3.6) - (3.8). Расчетные значения взяты из табл. 3.1.

Расчетные значения критерия равны

$$t_{b_0} = \left| \frac{b_0}{m_{b_0}} \right| = 11,45, \quad t_{b_1} = \left| \frac{b_1}{m_{b_1}} \right| = 18,66.$$

Табличное значение t-критерия Стьюдента при уровне значимости $\alpha = 0,05$ и числе степеней свободы $df = n - m - 1 = 36 - 1 - 1 = 34$ составляет $t_{\text{табл}}(0,05; 34) = 2,03$.

Поскольку $t_{b_0} > t_{\text{табл}}$, параметр b_0 статистически значим; $t_{b_1} > t_{\text{табл}}$, параметр b_1 также статистически значим.

Уравнение тренда имеет хорошее качество, адекватно описывает исходный ряд.

Таблица 3.3. Расчетная таблица оценки модели ряда

t	y_t	\hat{y}_t	$\left \frac{y_t - \hat{y}_t}{y_t} \right $	ε_t	ε_{t-1}	ε_t^2	$(\varepsilon_t - \varepsilon_{t-1})^2$
1	727,3	817,3	0,124	-90,0		8098,84	
2	868,7	875,6	0,008	-6,9	-90,0	47,61	6904,57
3	965,2	933,9	0,032	31,3	-6,9	979,31	1458,76
4	1041,2	992,2	0,047	49,0	31,3	2399,80	313,07
5	954,3	1050,5	0,101	-96,2	49,0	9258,00	21084,85
6	1058,5	1108,8	0,048	-50,3	-96,2	2532,58	2106,24
7	1150,0	1167,1	0,015	-17,1	-50,3	293,47	1101,83
8	1072,0	1225,4	0,143	-153,4	-17,1	23542,99	18579,39
9	1053,0	1283,7	0,219	-230,7	-153,4	53242,55	5976,26
10	1539,9	1342,0	0,128	197,9	-230,7	39144,73	183692,61
11	1507,3	1400,4	0,071	106,9	197,9	11437,03	8263,94

t	y_t	\hat{y}_t	$\left \frac{y_t - \hat{y}_t}{y_t} \right $	ε_t	ε_{t-1}	ε_t^2	$(\varepsilon_t - \varepsilon_{t-1})^2$
12	2013,0	1458,7	0,275	554,3	106,9	307290,39	200161,17
13	1345,3	1517,0	0,128	-171,7	554,3	29470,06	527085,06
14	1587,0	1575,3	0,007	11,7	-171,7	137,48	33633,27
15	1655,2	1633,6	0,013	21,6	11,7	467,38	97,89
16	1723,0	1691,9	0,018	31,1	21,6	968,01	90,13
17	1701,5	1750,2	0,029	-48,7	31,1	2371,05	6369,04
18	1923,9	1808,5	0,060	115,4	-48,7	13317,24	26926,76
19	1852,6	1866,8	0,008	-14,2	115,4	201,81	16797,78
20	1987,6	1925,1	0,031	62,5	-14,2	3904,73	5881,93
21	2003,3	1983,4	0,010	19,9	62,5	395,28	1815,29
22	1971,8	2041,7	0,035	-69,9	19,9	4889,45	8065,16
23	2027,0	2100,0	0,036	-73,0	-69,9	5333,51	9,65
24	2492,7	2158,3	0,134	334,4	-73,0	111798,54	165969,67
25	1793,8	2216,6	0,236	-422,8	334,4	178796,50	573361,29
26	2146,9	2274,9	0,060	-128,0	-422,8	16396,70	86903,36
27	2422,3	2333,3	0,037	89,0	-128,0	7928,86	47129,70
28	2380,8	2391,6	0,005	-10,8	89,0	115,82	9961,29
29	2315,4	2449,9	0,058	-134,5	-10,8	18081,73	15303,23
30	2632,8	2508,2	0,047	124,6	-134,5	15531,50	67129,58
31	2584,1	2566,5	0,007	17,6	124,6	310,44	11450,34
32	2606,7	2624,8	0,007	-18,1	17,6	327,14	1274,94
33	2547,5	2683,1	0,053	-135,6	-18,1	18385,54	13807,72
34	2485,6	2741,4	0,103	-255,8	-135,6	65433,40	14449,54
35	2611,8	2799,7	0,072	-187,9	-255,8	35308,58	4609,56
36	3406,5	2858,0	0,161	548,5	-187,9	300839,07	542275,77
Сумма	66155,5	66155,5	2,566	0	-548,5	1288977,1	2630040,6
Среднее	1837,7	1837,7	0,071	0			

3.4. Анализ отклонений от тренда

Согласно аддитивной модели временного ряда (1.1)

$$y_t = \hat{y}_t + \varepsilon_t$$

остатки (отклонения от тренда) ε_t находятся как разность между фактическими (наблюдаемыми) и теоретическими значениями уровней ряда:

$$\varepsilon_t = y_t - \hat{y}_t \quad (3.9)$$

В линейной модели остатки должны удовлетворять требованиям теоремы Гаусса-Маркова.

1. Проверка требования $M(\varepsilon_t) = 0$.

Числовой оценкой математического ожидания $M(\varepsilon_t)$ является среднее значение $M(\varepsilon_t) = \bar{\varepsilon}_t$. Для выполнения требования $M(\varepsilon_t) = 0$ необходимо, чтобы

$$\bar{\varepsilon}_t = \frac{1}{n} \sum_{t=1}^n \varepsilon_t \approx 0 \quad (3.10)$$

Дисперсия остатков ε_t должна быть одинаковой для всех значений t (свойство *гомоскедастичности*). Если это условие не соблюдается, имеет место *гетероскедастичность* остатков.

Эффективным методом предварительного анализа однородности остатков является визуальный анализ графиков остатков. Как правило, рассматривается график остатков в зависимости от номера уровня ряда (рис. 12, 13).

В случае гомоскедастичности точки (остатки) равномерно располагаются внутри горизонтальной полосы, симметричной оси абсцисс – оси времени t (рис. 12).

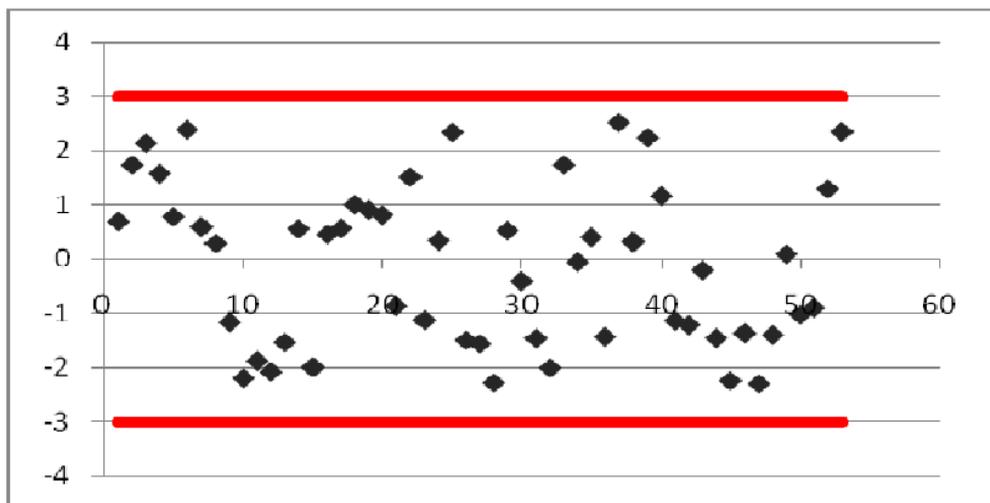


Рис. 12. Гомоскедастичность остатков

В случае гетероскедастичности точки (остатки), начиная с некоторого момента времени, устойчиво выходят из горизонтальной полосы, симметричной оси абсцисс (рис. 13).

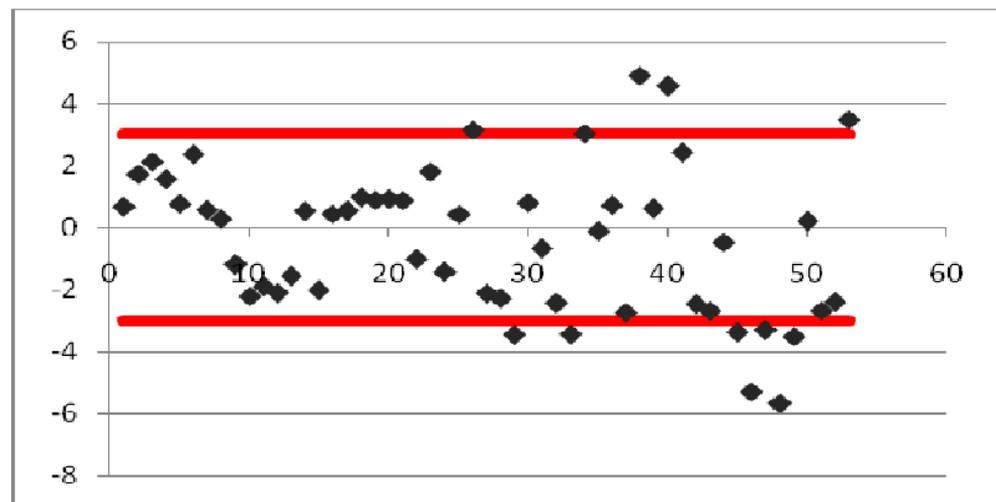


Рис. 13. Гетероскедастичность остатков

Для обнаружения эффекта гетероскедастичности, помимо визуального анализа, используют тесты Уайта, Голдфелда-Квандта, Глейзера и др.

Точечная (числовая) оценка дисперсии остатков σ_ε^2 находится по формуле:

$$s_\varepsilon^2 = \frac{1}{(n-2)} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \approx \sigma_\varepsilon^2 \quad (3.11)$$

2. Проверка требования: для разных наблюдений остатки ε_t некоррелированы (независимы).

Наиболее распространенный метод проверки требования о независимости остатков - критерий Дарбина-Уотсона (о наличии в остатках автокорреляции первого порядка), в котором рассчитывается статистика

$$d_{расч} = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2}, \quad (3.12)$$

где $0 \leq d_{расч} \leq 4$.

Для линейного уравнения тренда теоретические значения критерия Дарбина-Уотсона d_U и d_L находятся по таблице критических значений (табл. Приложения 4) по объему выборки n , числу степеней свободы $df=1$ и уровню значимости α .

С помощью критических значений числовой промежуток $(0; 4)$ разбивается на пять отрезков и принимается решение о независимости остатков.

Принятие решения о зависимости остатков

Есть положительная автокорреляция остатков	Зона неопределенности	Автокорреляция остатков отсутствует	Зона неопределенности	Есть отрицательная автокорреляция остатков		
0	d_L	d_U	2	$4 - d_U$	$4 - d_L$	4

На практике в случае, когда статистика критерия $d_{расч}$ попадает в зону неопределенности, обычно признается наличие автокорреляции в остатках.

Замечание

Если уравнение линейного тренда признано качественным, а остатки удовлетворяют требованиям теоремы Гаусса-Маркова, то аддитивная модель временного ряда считается качественной, т.е. она адекватно описывает исходный ряд.

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. построен линейный тренд. Провести анализ отклонений от тренда.

Проверим требования теоремы Гаусса–Маркова (расчеты см. в табл. 3.3) к остаткам модели временного ряда (3.9).

$$\varepsilon_t = y_t - \hat{y}_t$$

1. Среднее значение остатков (3.10) $\bar{\varepsilon} = \frac{1}{n} \sum_{t=1}^n \varepsilon_t = 0$.

Требование равенства нулю математического ожидания остатков выполняется.

2. По виду графика (рис. 14) можно сделать вывод о гетероскедастичности остатков (точки графика не находятся внутри горизонтальной полосы, симметричной относительно оси абсцисс). Требование постоянства дисперсии остатков не выполняется.

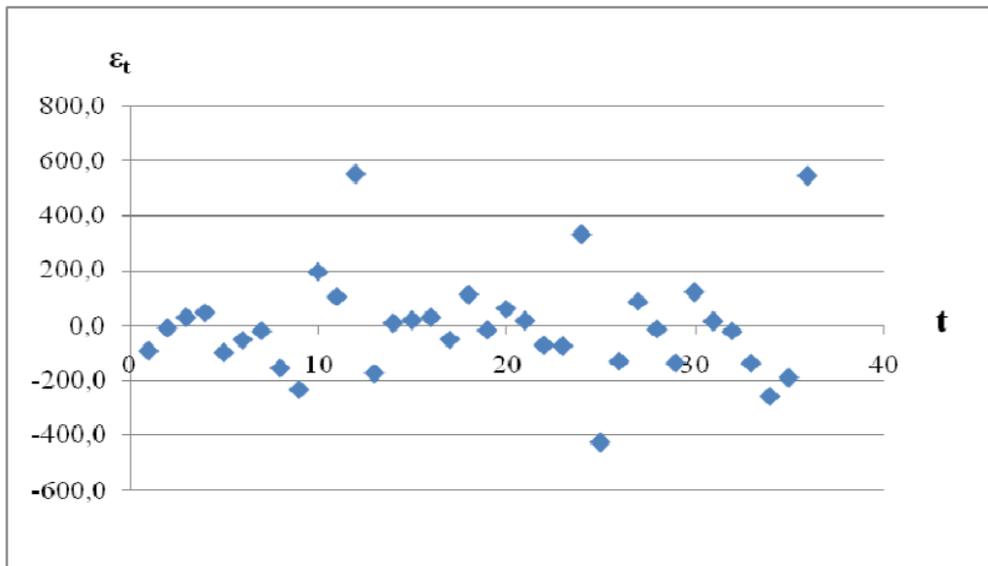


Рис. 14. График остатков

Дисперсия остатков равна (3.11)

$$s_{\varepsilon}^2 = \frac{1}{(n-2)} \sum_{t=1}^n (y_t - \hat{y}_t)^2 = \frac{1}{36-2} 1288977,1 = 37911,1 \neq \text{const}$$

3. Проверим наличие автокорреляции в остатках с помощью критерия Дарбина–Уотсона.

Найдем статистику (3.12) (табл. 3.3)

$$d_{\text{расч}} = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2} = \frac{2630040,6}{1288977,1} = 2,04.$$

По таблице критических точек распределения Дарбина–Уотсона по уровню значимости $\alpha = 0,05$, числу независимых переменных в уравнении тренда $m = 1$ и длине ряда $n = 36$ найдем критические значения $d_L = 1,40$ и $d_U = 1,52$.

Рассчитаем значения $4 - d_U = 2,48$, $4 - d_L = 2,60$.

С помощью критических значений числовой промежутки (0; 4) разобьем на пять отрезков:

I	I	I	I	I	I
0	1,40	1,52	2,48	2,60	4

В нашем случае, статистика критерия $d_{\text{расч}} = 2,04$ попадает в промежуток от $d_U = 1,52$ до $4 - d_U = 2,48$, следовательно, автокорреляция в остатках отсутствует. Требование независимости остатков выполняется.

3.5. Модель ряда и ее характеристики

Для оценки тесноты зависимости уровней ряда y_t от времени t используют коэффициент парной корреляции и шкалу Чеддока (табл. 2.4)

$$r_t = \frac{\sum_{t=1}^n (t - \bar{t}) \cdot (y_t - \bar{y}_t)}{\sqrt{\sum_{t=1}^n (t - \bar{t})^2 \cdot \sum_{t=1}^n (y_t - \bar{y}_t)^2}} \quad (3.13)$$

Для оценки доли вариации уровней ряда y_t , обусловленной временем t , используется коэффициент детерминации R^2 :

$$R^2 = (r_{yx})^2 \quad (0 \leq R^2 \leq 1) \quad (3.14)$$

Результаты исследований целесообразно свести в таблицу, которую можно рассматривать как модель временного ряда.

Аддитивная модель

$$y_t = \hat{y}_t + \varepsilon_t, \alpha =$$

Тренд	Отклонения от тренда
Линейный $\hat{y}_t = b_0 + b_1 \cdot t$	$\varepsilon_t = y_t - \hat{y}_t$
$\bar{A} =$	$\bar{\varepsilon}_t =$
$F_{расч} =$	$s_{\varepsilon}^2 \approx \sigma_{\varepsilon}^2 = (\neq) const$
$F_{крит} =$	$d_{расч} =$
$t_{b_0} =$	$d_L =$
$t_{b_1} =$	$d_U =$
$t_{крит} =$	
$r_t =$	
$R^2 =$	

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. построен линейный тренд. Записать аддитивную модель ряда и ее характеристики.

Определим тесноту связи уровней ряда со временем, т.е. со всей совокупностью взаимосвязанных внешних факторов, и их влияние на временной ряд, рассчитав коэффициент корреляции (3.13) $r_t = 0,95$ и коэффициент детерминации (3.14) $R^2 = 0,911$.

Запишем модель исследуемого временного ряда в виде таблицы.

Аддитивная модель $y_t = \hat{y}_t + \varepsilon_t, \alpha = 0,05$	
Тренд	Отклонения от тренда
Линейный $\hat{y}_t = 758,99 + 58,31 \cdot t$ (t=1, 2,..., 36)	$\varepsilon_t = y_t - \hat{y}_t$
$\bar{A} = 7,13\%$	$\bar{\varepsilon}_t \approx 0$
$F_{расч} = 348,38$	$s_{\varepsilon}^2 \approx \sigma_{\varepsilon}^2 = 37911,1 \neq const$
$F_{крит} = 4,13$	$d_{расч} = 2,04$
$t_{b0} = 11,45$	$d_L = 1,40$
$t_{b1} = 18,66$	$d_U = 1,52$
$t_{крит} = 2,03$	
$r_t = 0,96$	
$R^2 = 0,911$	

Параметр уравнения тренда $b_0 = 758,99$ интерпретируется как параметр начальных условий (значение переменной y_t в нулевой момент времени, т.е. денежных доходов в момент $t=0$, что соответствует трендовому значению доходов в декабре 1998 г.). Параметр $b_1 = 58,31$ – скорость роста переменной y_t (т.е. в среднем за 1999-2001 г. доходы населения увеличиваются на 58,31 руб. в месяц).

Коэффициент парной корреляции $r_t = 0,96 > 0,7$ (табл. 3.1) показывает, что доходы населения сильно

зависят от времени (внешних факторов) и на 91,1% объясняются временем (внешними факторами) (коэффициент детерминации $R^2 = 0,911$).

Построенная модель временного ряда $y_t = 758,99 + 58,31 \cdot t + \varepsilon_t$ имеет высокую точность ($\bar{A} = 7,13\%$), на уровне 5%-й ошибки статистически значима в целом, параметры тренда также статистически значимы. Отклонения от тренда имеют нулевое среднее и независимы. В целом модель обладает хорошим качеством и может быть использована для прогнозирования. Однако она имеет один недостаток – непостоянство дисперсии остатков, который объясняется тем, что в модели не учитываются циклические колебания, имеющие наибольшую величину отклонения от тренда декабрьских и январских значений доходов. При прогнозировании по данной модели прогнозные значения будущих декабрьских и январских доходов будут наименее точными.

3.6. Моделирование временного ряда в условиях смены тенденции. Тест Чоу

При протекании экономических процессов иногда возникает ситуация единовременного изменения характера динамики изучаемого показателя. В этом случае, начиная с некоторого момента t^* , происходит изменение тенденции временного ряда. Момент времени t^* обычно называют точкой смены тенденции.

Основная задача исследования временного ряда, включающего точку смены тенденции, – выяснить насколько значимо повлияли структурные изменения на

характер тенденции. Если это влияние значимо, для моделирования тенденции необходимо разделить ряд на две части (до момента времени t^* и после момента t^*) и построить отдельно по каждой части ряда уравнения тренда (построить «кусочную» модель). Если изменения незначительно повлияли на характер тенденции, то ее можно описать единым для всего ряда уравнением.

Например, на рис. 15 изображен график ряда, в котором произошла смена тенденции в промежуток времени между $t = 28$ и $t = 34$, близкий к моменту времени $t^* = 32$.

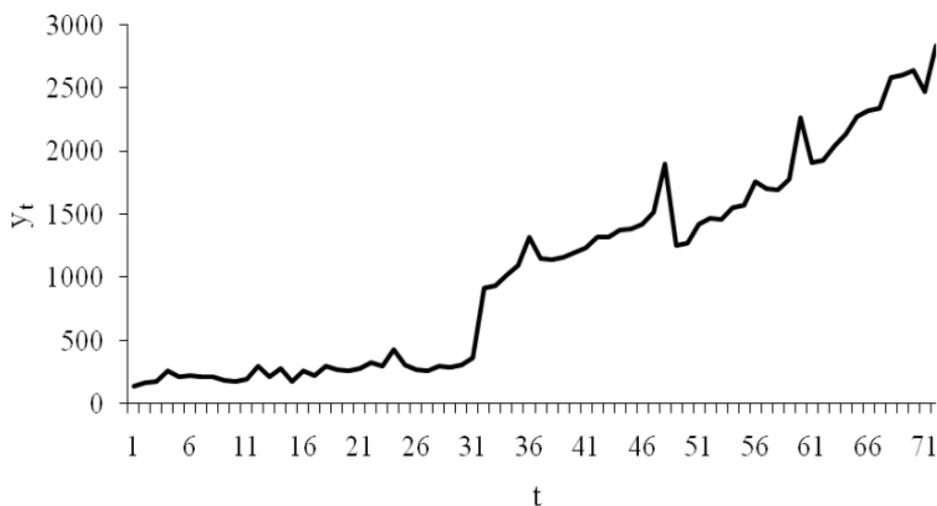


Рис. 15. Смена тенденции временного ряда.

Значимость структурных изменений можно оценить с помощью статистического критерия Грегори Чоу. Система обозначений для его проверки приведена в табл. 3.4.

Таблица 3.4. Система обозначений теста Чоу

№ уравнения	Вид уравнения тренда*	Длина ряда	Остаточная сумма квадратов	Число параметров в уравнении тренда	Число степеней свободы остаточной дисперсии
Кусочно-линейная модель ряда					
(I)	$y^{(1)} = a_1 + b_1 \cdot t$	n_1	$S^1_{\text{ост}}$	k_1	$n_1 - k_1$
(II)	$y^{(2)} = a_2 + b_2 \cdot t$	n_2	$S^2_{\text{ост}}$	k_2	$n_2 - k_2$
Уравнение тренда по всему ряду					
(III)	$y^{(3)} = a_3 + b_3 \cdot t$	$n = n_1 + n_2$	$S^3_{\text{ост}}$	k_3	$n - k_3 = (n_1 + n_2) - k_3$
* В общем случае вид уравнений трендов и число параметров в каждом уравнении могут различаться					

Выдвинем гипотезу H_0 о стабильности тенденции изучаемого временного ряда y_t , при конкурирующей гипотезе H_1 о наличии нарушения тенденции в точке t^* .

Найдем остаточные суммы квадратов

$$S^1_{\text{ост}} = \sum_{t=1}^{n_1} (y_t - \hat{y}_t^{(1)})^2, \quad (3.15)$$

$$S^2_{\text{ост}} = \sum_{t=n_1+1}^n (y_t - \hat{y}_t^{(2)})^2, \quad (3.16)$$

$$S^3_{\text{ост}} = \sum_{t=1}^n (y_t - \hat{y}_t^{(3)})^2, \quad (3.17)$$

где $\hat{y}_t^{(1)}$, $\hat{y}_t^{(2)}$, $\hat{y}_t^{(3)}$ - теоретические значения уровней, найденные соответственно по уравнениям (I), (II) и (III) (табл. 3.4).

Остаточная сумма квадратов кусочно-линейной модели равна

$$S_{ост}^{кл} = S_{ост}^1 + S_{ост}^2 . \quad (3.18)$$

Соответствующее ей число степеней свободы составляет

$$(n_1 - k_1) + (n_2 - k_2) = (n - k_1 - k_2) .$$

Изменение остаточной дисперсии при переходе от единого уравнения тренда к кусочно-линейной модели определяется как разность

$$\Delta S_{ост} = S_{ост}^3 - S_{ост}^{кл} ,$$

соответствующее число степеней свободы равно

$$n - k_3 - (n - k_1 - k_2) = k_1 + k_2 - k_3 .$$

Расчетное значение Теста Чоу

$$F_{расч} = \frac{\Delta S_{ост}}{S_{ост}^{кл}} \cdot \frac{n - k_1 - k_2}{k_1 + k_2 - k_3} . \quad (3.19)$$

сравнивается с табличным $F_{табл}(\alpha, df_1, df_2)$, найденным по таблице критических точек распределения Фишера-

Снедекора для уровня значимости α и числа степеней свободы $df_1 = k_1 + k_2 - k_3$, $df_2 = n - k_1 - k_2$, или с помощью встроенной функции Excel «ФРАСПОБР»

Если $F_{\text{расч}} > F_{\text{табл}}$, то гипотеза H_0 о структурной стабильности отклоняется, принимается гипотеза H_1 , влияние структурных изменений на динамику изучаемого показателя считается значимым. Момент времени t^* является точкой смены тенденции. В этом случае моделирование тенденции временного ряда необходимо проводить с помощью кусочной модели.

В противном случае, когда $F_{\text{расч}} < F_{\text{табл}}$, нет оснований отклонять гипотезу H_0 , и моделирование тенденции следует осуществлять с помощью единого для всего ряда уравнения тренда.

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1996 г. по декабрь 2001 г. проверить гипотезу о смене тенденции с помощью критерия Грегори Чоу (взять значение $t^* = 32$).

Значимость структурных изменений в формировании денежных доходов населения оценим с помощью критерия Грегори Чоу.

Разделим исходный ряд на две части (в первой – 32 первых уровней исходного ряда, во второй – последние 40 уровней). Выбор предполагаемой точки смены тенденции ($t^* = 32$) осуществлялся по виду графика (рис. 7).

Для проверки теста Чоу построим уравнения линейных трендов для каждой части ряда и по ним найдем

остаточные суммы квадратов. Расчеты рекомендуется проводить с помощью пакета Excel «Анализ данных», инструмент «Регрессия». Результаты расчетов с помощью пакета приведены в табл. 3.5, 3.6, 3.7, в которых значения параметров уравнений трендов и остаточных сумм квадратов выделены жирным шрифтом.

Таблица 3.5. Результаты регрессионного анализа для первой части ряда

ВЫВОД ИТОГОВ

Регрессионная статистика

Множественный R	0,814
R-квадрат	0,663
Нормированный R-квадрат	0,652
Стандартная ошибка	56,058
Наблюдения	32

Дисперсионный анализ

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	1	185507,06	185507,06	59,03	0,000
Остаток	30	94275,87	3142,53		
Итого	31	279782,94			

	<i>Кэф-фициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	423,52	20,29	20,87	0,000	382,07	464,96
t	8,25	1,07	7,68	0,000	6,05	10,44

Уравнение тренда для первой части ряда имеет вид

$$\hat{y}_t = 423,52 + 8,25 \cdot t, \quad (t = 1, 2, \dots, 32)$$

Таблица 3.6. Результаты регрессионного анализа для второй части ряда

ВЫВОД ИТОГОВ

Регрессионная статистика

Множественный R	0,962
R-квадрат	0,925
Нормированный R-квадрат	0,923
Стандартная ошибка	190,732
Наблюдения	40

Дисперсионный анализ

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	1	17125913,11	17125913,11	470,77	0,00
Остаток	38	1382392,25	36378,74		
Итого	39	18508305,36			

	<i>Кэф-фициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	-1242,58	140,43	-8,85	0,000	-1526,87	-958,28
t	56,68	2,61	21,70	0,000	51,40	61,97

Уравнение тренда для второй части ряда имеет вид

$$\hat{y}_t = -1242,58 + 56,68 \cdot t, \quad (t = 33, 34, \dots, 72)$$

Таблица 3.7. Результаты регрессионного анализа для всего ряда

ВЫВОД ИТОГОВ

Регрессионная статистика

Множественный R	0,927
R-квадрат	0,860
Нормированный R-квадрат	0,858
Стандартная ошибка	294,293
Наблюдения	72

Дисперсионный анализ

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	1	37218784,78	37218784,78	429,74	0,00
Остаток	70	6062570,78	86608,15		
Итого	71	43281355,56			

	<i>Кэф-фициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	-51,05	70,09	-0,73	0,47	-190,85	88,75
t	34,60	1,67	20,73	0,000	31,27	37,92

Уравнение тренда для всего ряда имеет вид

$$\hat{y}_t = -51,05 + 36,60 \cdot t, \quad (t = 1, 2, \dots, 72)$$

Остаточные суммы квадратов (ESS) находятся по формулам (3.15) - (3.18):

Результаты расчетов сведены в табл. 3.8.

Таблица 3.8. Расчетная таблица теста Чоу

	Длина ряда n	Остаточная сумма квадратов $S_{ост}^2$	Число параметров в уравнении тренда k
Первая часть ряда (1)	32	94275,872	2
Вторая часть ряда (2)	40	1382392,246	2
Весь ряд (3)	72	6062570,775	2

В рассматриваемом примере все уравнения тренда линейные, число параметров всех уравнений $k_1 = k_2 = k_3 = 2$.

Вычислим сумму

$$S_{ост}^2 = S_{ост}^2 (1) + S_{ост}^2 (2) = 94275,872 + 1382392,246 = 1476668,118.$$

Найдем разность

$$\Delta S_{ост}^2 = S_{ост}^2 (3) - S_{ост}^2 = 6062570,775 - 1476668,18 = 4585902,657.$$

Вычислим статистику критерия

$$F_{расч} = \frac{\Delta S_{ост}^2}{S_{ост}^2} \cdot \frac{n - k_1 - k_2}{k_1 + k_2 - k_3} =$$

$$= \frac{4585902,657}{1476668,118} \cdot \frac{72-2-2}{2+2-2} = 105,59.$$

Найденное значение $F_{расч}$ сравним с критическим значением $F_{табл}(\alpha; df_1; df_2)$, где

$$df_1 = k_1 + k_2 - k_3 = 2 + 2 - 2 = 2,$$

$$df_2 = n - k_1 - k_2 = 72 - 2 - 2 = 68.$$

Критическое (табличное) значение распределения Фишера

$$F_{табл}(0,05;2;68) = 3,13.$$

Поскольку в нашем случае $F_{расч} > F_{табл}$, структурные изменения во временном ряду считаются статистически значимыми. Точка $t^* = 32$ является точкой смены тенденции.

Моделирование тенденции временного ряда необходимо проводить с помощью кусочной модели (двух уравнений тренда, построенным по уровням ряда до и после точки смены тенденции).

Действительно, в августе 1998г. ($t^* = 32$) в России произошел экономический кризис (дефолт), приведший к значительному изменению тенденции роста денежных доходов населения. Однако для формирования новой тенденции необходимо время. Будем считать, что новая тенденция сформировалась к началу 1999 г. Далее в рассматриваемом примере анализ показателя «Среднедушевые денежные доходы» проводится с января 1999 г. (для уровней исходного ряда, начиная с $t = 37$).

4. Моделирование временного ряда с учетом сезонности

4.1. Аддитивная модель с учетом сезонности

Наиболее простым подходом к анализу временных рядов, содержащих сезонные или циклические колебания, является расчет значений сезонной компоненты методом скользящей средней и построение модели временного ряда.

Существует два типа моделей: аддитивная и мультипликативная.

Если амплитуда колебаний уровней ряда приблизительно постоянна (рис. 16), строят аддитивную модель вида (1.1):

$$y_t = \hat{y}_t + s_t + \varepsilon_t ,$$

где \hat{y}_t – трендовая, s_t – циклическая (сезонная), ε_t – случайная составляющие (компоненты) уровней ряда.

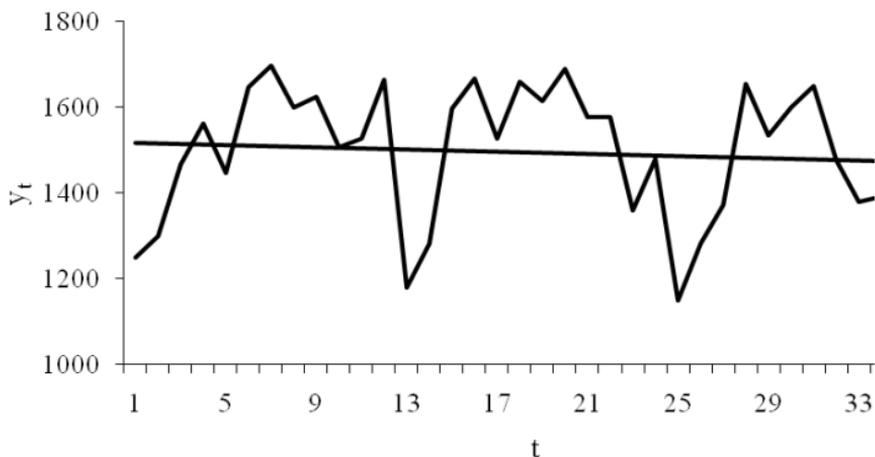


Рис. 16. Аддитивные циклические колебания

Если амплитуда колебаний уровней во времени растет (рис. 17) (уменьшается), строят мультипликативную модель вида (1.2):

$$y_t = \hat{y}_t \times s_t \times \varepsilon_t .$$

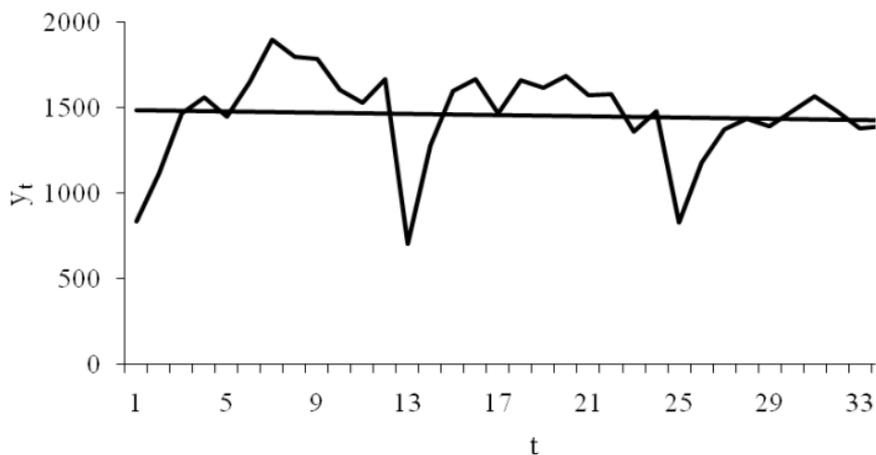


Рис. 17. Мультипликативные циклические колебания

Построение этих моделей сводится к расчету значений \hat{y}_t , s_t , ε_t для каждого уровня ряда и производится в следующем порядке (алгоритме).

1. Производится выравнивание исходного ряда методом скользящей средней.

2. Рассчитываются значения сезонной компоненты s_t . Необходимо помнить, что в аддитивной модели сумма скорректированных сезонных компонент внутри цикла равна нулю. В мультипликативной модели произведение скорректированных сезонных компонент равно длине цикла.

3. Устраняются скорректированные сезонные компоненты S_t из исходных уровней ряда. Для аддитивной модели рассчитывается разность $(y_t - s_t)$, для мультипликативной – частное (y_t / s_t) . Затем получают выравненные данные: $(\hat{y}_t + \varepsilon_t)$ - для аддитивной, или $(\hat{y}_t \cdot \varepsilon_t)$ - для мультипликативной модели.

4. Производится аналитическое выравнивание уровней (т.е. находится уравнение тренда) $(\hat{y}_t + \varepsilon_t)$ или $(\hat{y}_t \cdot \varepsilon_t)$ и рассчитываются теоретические значения \hat{y}_t по полученному уравнению тренда.

5. Рассчитываются по модели значения $(\hat{y}_t + s_t)$ или $(\hat{y}_t \cdot s_t)$.

6. Находятся случайные ошибки ε_t .

Пример.

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. построить аддитивную модель временного ряда с учетом сезонности.

Построим аддитивную модель доходов населения с линейным трендом по указанному выше алгоритму. Расчеты сведены в табл. 4.1.

Таблица 4.1. Построение аддитивной модели временного ряда с сезонной составляющей

t	y_t	\tilde{y}_t	S_t	S_t^{ck}	$y_t - S_t^{ck}$	\hat{y}_t	ε_t
1	2	3	4	5	6	7	8
1	727,30	786,04	-58,74	-236,44	963,74	850,05	113,70
2	868,70	840,00	28,70	-47,00	915,70	906,48	9,22
3	965,20	911,34	53,86	108,51	856,69	962,91	-106,23
4	1041,20	977,58	63,62	26,02	1015,18	1019,35	-4,17
5	954,30	1033,84	-79,54	-97,52	1051,82	1075,78	-23,97
6	1058,50	1055,20	3,30	75,58	982,92	1132,22	-149,30
7	1150,00	1057,56	92,44	35,49	1114,51	1188,65	-74,14
8	1072,00	1174,68	-102,68	-6,38	1078,38	1245,09	-166,71
9	1053,00	1264,44	-211,44	-62,61	1115,61	1301,52	-185,91
10	1539,90	1437,04	102,86	-86,48	1626,38	1357,96	268,42
11	1507,30	1491,70	15,60	-99,61	1606,91	1414,39	192,52
12	2013,00	1598,50	414,50	390,42	1622,58	1470,83	151,75
13	1345,30	1621,56	-276,26	-236,44	1581,74	1527,26	54,48
14	1587,00	1664,70	-77,70	-47,00	1634,00	1583,70	50,30
15	1655,20	1602,40	52,80	108,51	1546,69	1640,13	-93,44
16	1723,00	1718,12	4,88	26,02	1696,98	1696,57	0,41
17	1701,50	1771,24	-69,74	-97,52	1799,02	1753,00	46,01
18	1923,90	1837,72	86,18	75,58	1848,32	1809,44	38,89
19	1852,60	1893,78	-41,18	35,49	1817,11	1865,87	-48,76
20	1987,60	1947,84	39,76	-6,38	1993,98	1922,30	71,68
21	2003,30	1968,46	34,84	-62,61	2065,91	1978,74	87,17
22	1971,80	2096,48	-124,68	-86,48	2058,28	2035,17	23,10
23	2027,00	2057,72	-30,72	-99,61	2126,61	2091,61	35,00
24	2492,70	2086,44	406,26	390,42	2102,28	2148,04	-45,77
25	1793,80	2176,54	-382,74	-236,44	2030,24	2204,48	-174,24
26	2146,90	2247,30	-100,40	-47,00	2193,90	2260,91	-67,02

t	y_t	\tilde{y}_t	S_t	S_t^{ck}	$y_t - S_t^{ck}$	\hat{y}_t	ε_t
1	2	3	4	5	6	7	8
27	2422,30	2211,84	210,46	108,51	2313,79	2317,35	-3,56
28	2380,80	2379,64	1,16	26,02	2354,78	2373,78	-19,01
29	2315,40	2467,08	-151,68	-97,52	2412,92	2430,22	-17,30
30	2632,80	2503,96	128,84	75,58	2557,22	2486,65	70,57
31	2584,10	2537,30	46,80	35,49	2548,61	2543,09	5,52
32	2606,70	2571,34	35,36	-6,38	2613,08	2599,52	13,56
33	2547,50	2567,14	-19,64	-62,61	2610,11	2655,96	-45,85
34	2485,60	2731,62	-246,02	-86,48	2572,08	2712,39	-140,32
35	2611,80	2903,93	-292,13	-99,61	2711,41	2768,83	-57,41
36	3406,50	3064,40	342,10	390,42	3016,08	2825,26	190,81
Сумма			-100,97	0,00			
Среднее	1837,65	1840,46			1837,65	1837,65	0

В 3-м столбце таблицы записаны результаты численного сглаживания по $m = 5$ точкам по формулам (2.5), (2.6) (пункт 1 алгоритма).

В 4-м столбце таблицы записаны сезонные составляющие (пункт 2 алгоритма).

В 5-м столбце таблицы записаны скорректированные сезонные составляющие (пункт 2 алгоритма).

Корректировка сезонных составляющих производится в табл. 4.2.

Таблица 4.2. Расчет скорректированных значений сезонной компоненты

Номер месяца	Год			\bar{S}_t	S_t^{ck}
	1999	2000	2001		
1	-58,74	-276,26	-382,74	-239,25	-236,44
2	28,70	-77,70	-100,40	-49,80	-47,00
3	53,86	52,80	210,46	105,71	108,51
4	63,62	4,88	1,16	23,22	26,02
5	-79,54	-69,74	-151,68	-100,32	-97,52
6	3,30	86,18	128,84	72,77	75,58
7	92,44	-41,18	46,80	32,69	35,49
8	-102,68	39,76	35,36	-9,19	-6,38
9	-211,44	34,84	-19,64	-65,41	-62,61
10	102,86	-124,68	-246,02	-89,28	-86,48
11	15,60	-30,72	-292,13	-102,42	-99,61
12	414,50	406,26	342,10	387,62	390,42
Сумма				-33,66	0,00

Корректировочный коэффициент $k = \frac{-33,66}{12} = -2,80$.

В 6-м столбце таблицы скорректированные сезонные компоненты устранены из уровней исходного ряда (пункт 3 алгоритма).

По данным 6-го столбца таблицы находится уравнение линейного тренда с помощью пакета «Анализ данных» инструмент «Регрессия» (пункт 4 алгоритма) (см. табл. 4.3). Теоретические значения уровней ряда, найденные по уравнению тренда (3.1) записаны в 7-м столбце.

Таблица 4.3. Результаты расчетов параметров и характеристик линейного тренда в аддитивной модели с сезонной составляющей

ВЫВОД ИТОГОВ

Регрессионная статистика

Множественный R	0,985
R-квадрат	0,970
Нормированный R-квадрат	0,969
Стандартная ошибка	106,579
Наблюдения	36

Дисперсионный анализ

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	1	12373252,70	12373252,70	1089,28	0,00
Остаток	34	386210,50	11359,13		
Итого	35	12759463,20			

	<i>Кoeffициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	793,61	36,28	21,87	0,00	719,88	867,34
t	56,43	1,71	33,00	0,00	52,96	59,91

Уравнение тренда имеет вид $\hat{y}_t = 793,61 + 56,43 \cdot t$
 ($t = 1, 2, \dots, 36$) (рис. 18).

В 8-м столбце вычислены случайные компоненты (пункт 6 алгоритма).

Точность модели определим с помощью средней относительной ошибки аппроксимации. Модельная

ошибка $\bar{A} = 5,23\% < 10\%$, следовательно, точность модели высокая. Расчеты приведены в табл. 4.4.

В табл. 4.4 модельное значение уровней ряда найдено по формуле

$$\hat{y}_{t, модель} = \hat{y}_t + S_t^{ck}. \quad (4.1)$$

Таблица 4.4. Расчет модельной ошибки аддитивной модели с сезонностью

t	y_t	$\hat{y}_{t, модель}$	$\left \frac{y_t - \hat{y}_{t, модель}}{y_t} \right $
1	727,30	613,60	0,1563
2	868,70	859,48	0,0106
3	965,20	1071,43	0,1101
4	1041,20	1045,37	0,0040
5	954,30	978,27	0,0251
6	1058,50	1207,80	0,1410
7	1150,00	1224,14	0,0645
8	1072,00	1238,71	0,1555
9	1053,00	1238,91	0,1766
10	1539,90	1271,48	0,1743
11	1507,30	1314,78	0,1277
12	2013,00	1861,25	0,0754
13	1345,30	1290,82	0,0405
14	1587,00	1536,70	0,0317
15	1655,20	1748,64	0,0565
16	1723,00	1722,59	0,0002
17	1701,50	1655,49	0,0270
18	1923,90	1885,01	0,0202
19	1852,60	1901,36	0,0263

t	y_t	$\hat{y}_{t, модель}$	$\frac{ y_t - \hat{y}_{t, модель} }{y_t}$
20	1987,60	1915,92	0,0361
21	2003,30	1916,13	0,0435
22	1971,80	1948,70	0,0117
23	2027,00	1992,00	0,0173
24	2492,70	2538,47	0,0184
25	1793,80	1968,04	0,0971
26	2146,90	2213,92	0,0312
27	2422,30	2425,86	0,0015
28	2380,80	2399,81	0,0080
29	2315,40	2332,70	0,0075
30	2632,80	2562,23	0,0268
31	2584,10	2578,58	0,0021
32	2606,70	2593,14	0,0052
33	2547,50	2593,35	0,0180
34	2485,60	2625,92	0,0565
35	2611,80	2669,21	0,0220
36	3406,50	3215,69	0,0560
			Сумма = 1,8824

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. построена аддитивная модель временного ряда с учетом сезонности. Записать аддитивную модель ряда и ее характеристики.

Запишем модель ряда в виде таблицы 4.5.

Таблица 4.5. Аддитивная модель временного ряда с сезонной составляющей и линейным трендом

Аддитивная модель $y_t = \hat{y}_t + s_t + \varepsilon_t, \alpha = 0,05$		
Тренд	Скорректированная сезонная компонента	
Линейный $\hat{y}_t = 793,61 + 56,43 \cdot t$ ($t = 1, 2, \dots, 36$)	Месяц	S_t^{ck}
$F_{расч} = 1089,28$	январь	-236,44
$F_{крит}(0,05, 1, 34) = 4,13$	февраль	-47,00
$t_{b_0} = 21,87$	март	108,51
$t_{b_1} = 33,00$	апрель	26,02
$t_{крит}(0,05, 34) = 2,03$	май	-97,52
$r_t = 0,99$	июнь	75,58
$R^2 = 0,970$	июль	35,49
	август	-6,38
	сентябрь	-62,61
	октябрь	-86,48
	ноябрь	-99,61
	декабрь	390,42
$\bar{A} = 5,23\%$		

Результаты моделирования изображены на рис. 18.

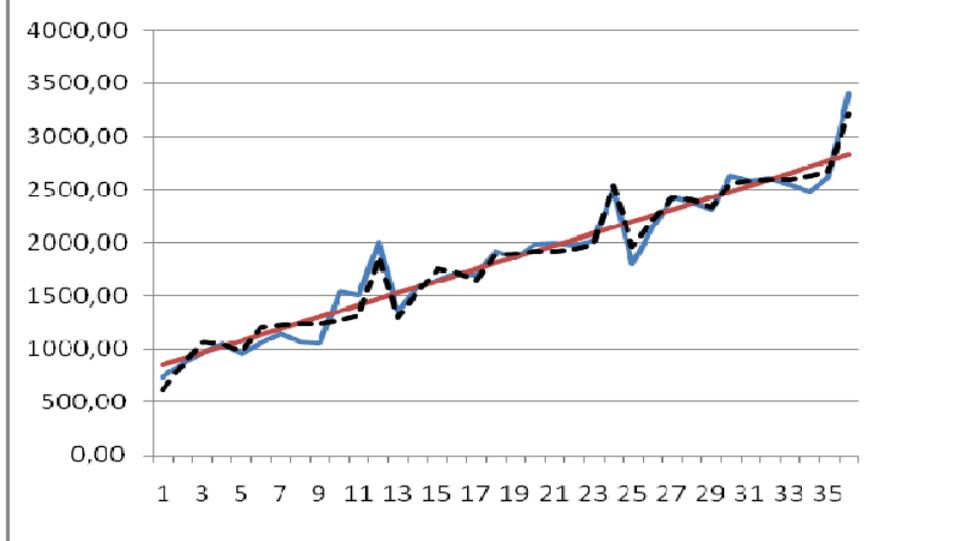


Рис. 18. График ряда, аддитивная модель с сезонностью и линия тренда

4.2. Определение степени полиномиального тренда

В случае, когда исследуемый экономический процесс носит колебательный характер, его динамику в некоторых случаях можно описать полиномиальным трендом (на рис. 19 динамика ряда описывается полиномом пятой степени).

Для аппроксимации ряда полиномом степени p предварительно находится значение этой степени по следующей процедуре, которая аналогична дифференцированию полинома. Очевидно, что вторая производная полинома первой степени, третья производная полинома второй степени и т.д. равны нулю. В случае временного ряда операция дифференцирования заменяется вычислением переменных разностей, а условие

равенства нулю - проверкой гипотезы о равенстве дисперсий предыдущих и последующих разностей.

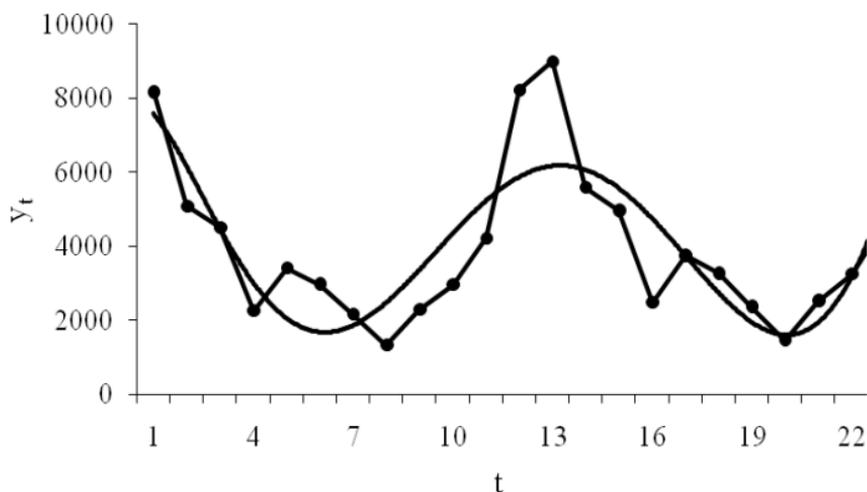


Рис. 19. Графики ряда и полиномиального тренда

Сначала вычисляются первые разности

$$\Delta^1 y_t = y_{t+1} - y_t, \text{ где } t = 1, 2, \dots, n - 1.$$

Затем по первым разностям вычисляются вторые разности

$$\Delta^2 y_t = \Delta^1 y_{t+1} - \Delta^1 y_t, \text{ где } t = 1, 2, \dots, n - 2.$$

И далее последовательно разности 3-го, ..., k-го порядков

$$\Delta^k y_t = \Delta^{k-1} y_{t+1} - \Delta^{k-1} y_t, \text{ где } t = 1, 2, \dots, n - k.$$

Под разностями нулевого порядка понимается сам временной ряд.

На каждом шаге, начиная с $k=0$, вычисляются дисперсии разностей k-го порядка по формуле

$$S_k^2 = \frac{\sum_{t=1}^{n-k} \left(\Delta^k y_t - \overline{\Delta^k y_t} \right)^2}{(n-k-1) \cdot (2k)!} \cdot (k!)^2. \quad (4.2)$$

При $k = 0$ вычисляется дисперсия заданного временного ряда y_t .

На каждом шаге с помощью критерия Фишера проверяется гипотеза о равенстве предыдущей и последующей дисперсий.

Для этого вычисляется расчетное значение критерия

$$F_{расч}^k = \frac{S_{k-1}^2}{S_k^2}. \quad (4.3)$$

Критическое значение критерия $F_{крит}^k(\alpha, df_1, df_2)$ находится в таблице критических значений распределения Фишера-Снедекора по уровню значимости α и числам степеней свободы $df_1 = n - k$, $df_2 = n - k - 1$.

Если $F_{расч}^k > F_{крит}^k$ дисперсии отличаются значимо.

В этом случае процедура вычислений дисперсий и их разностей продолжается. Доказано, что последовательность дисперсий (4.2) убывает с ростом k , и при некотором значении $p^* = k - 1$ выполнится неравенство $F_{расч}^k < F_{крит}^k$ (различие дисперсий становится незначимым). Полученное значение p^* и является степенью полиномиального тренда. Дисперсия s_k^2 называется дисперсией случайностей, а разности порядка p^* являются случайной компонентой временного ряда.

5. Методы прогнозирования временных рядов

Исследование динамики экономических явлений и процессов, выявление и характеристика основного тренда развития и моделей взаимосвязи дают основания для прогнозирования, т.е. определения будущих размеров уровней показателей, описывающих процессы.

Прогнозирование основывается на предположении, что закономерность развития, действующая в прошлом внутри ряда динамики, сохранится и в прогнозируемом будущем. Такой прогноз основан на перспективной интерполяции.

Теоретической основой распространения тенденции является инерционность социально-экономических процессов, которая позволяет выявить взаимосвязь между уровнями динамического ряда или между группой связанных временных рядов. Надежные результаты прогнозирования временных рядов получают, если уровни ряда динамики сопоставимы и синтезированы на основе единого методологического подхода.

Применение перспективной экстраполяции в практическом прогнозировании основывается на следующих предпосылках:

- 1) тенденция развития изучаемого явления (процесса) графически описывается плавной линией;
- 2) общая тенденция развития явления (процесса) в будущем существенно не меняется.

Надежность прогноза зависит от того, как точны эти предположения в действительности, а также как точно охарактеризована выявленная закономерность.

Экстраполяция - это начальная стадия построения прогноза. Механическое, без учета условий, предпосылок и содержательного эконометрического анализа, применение экстраполяции может стать причиной неадекватности выводов.

Чем шире временной горизонт прогноза (чем более долгосрочен прогноз), тем очевиднее недостаточность простого метода экстраполяции в результате изменения тенденции, влияния новых факторов и т.д. В этом случае динамичность экономических процессов противоречит инерционности их развития.

Так как исследуемые временные ряды часто имеют недостаточно большую длину n , то временной горизонт прогнозирования ограничен. Поэтому, чем короче срок прогнозирования (период упреждения), тем более надежны результаты прогноза. За относительно короткий период условия развития процесса не успевают измениться, что сохраняет характер его динамики.

В зависимости от принципов построения и эмпирических данных ряда выделяются следующие методы прогнозирования:

- 1) по среднему приросту;
- 2) по среднему темпу роста;
- 3) на основе численного сглаживания;
- 4) на основе аналитического выравнивания.

Существуют и другие, более сложные методы прогнозирования, которые не входят в содержание данного пособия.

5.1. Прогнозирование по среднему приросту

Прогнозирование по среднему приросту может быть выполнено в случае линейной тенденции развития. Этот метод основывается на предположении о стабильности (равномерности) изменении уровней ряда.

Для прогнозирования по среднему приросту необходимо определить средний прирост и последовательно увеличивать конечный уровень ряда на его величину на требуемое число периодов:

$$\hat{y}_{n+T} = y_n + \bar{\Delta} \cdot T, \quad (5.1)$$

где \hat{y}_{n+T} – прогнозируемый уровень ряда; T – срок прогноза (период упреждения); y_n – последний уровень ряда, за который рассчитан средний прирост $\bar{\Delta}$:

$$\bar{\Delta} = \frac{1}{n-1} \sum_{t=1}^{n-1} (y_{t+1} - y_t). \quad (5.2)$$

Следует иметь в виду, что использование среднего прироста для прогноза возможно только при условии

$$S_{ост}^2 \leq p^2,$$

где
$$p^2 = \frac{1}{2n} \sum_{t=2}^n (y_t - y_{t-1})^2,$$

$$S_{ост}^2 = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_{n+T})^2 .$$

В качестве оценки точности прогноза можно взять модельную ошибку, которая рассчитывается как средняя относительная ошибка аппроксимации

$$\bar{A} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_{t, модель}}{y_t} \right| \cdot 100\% ,$$

где y_t - уровни ряда, $\hat{y}_{t, модель}$ - уровни ряда, найденные по модели

$$\hat{y}_{t, модель} = y_0 + \bar{\Delta} \cdot (t - 1), (t = 2, 3, \dots, n) . \quad (5.3)$$

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. найти прогноз доходов населения в первом квартале 2002 г. по среднему приросту. Оценить точность прогноза с помощью модельной ошибки. Результаты анализа изобразить графически

Найдем прогноз доходов населения в первом квартале 2002 г. по среднему приросту. Все необходимые расчеты приведены в табл. 5.1.

Приросты найдем по формуле

$$\Delta_t = y_t - y_{t-1} , \quad (t = 2, 3, \dots, n) .$$

Средний прирост равен (5.2):

$$\bar{\Delta} = \frac{1}{n-1} \sum_{t=2}^n \Delta_t = 76,55.$$

Модельная ошибка равна

$$\bar{A} = \frac{1}{36} 0,0781 \cdot 100\% = 14,15\% .$$

Прогнозная модель имеет вид

$$\hat{y}_{n+T} = y_n + 76,55 \cdot T .$$

В январе 2002 г. ($T = 1$, $n + T = 36 + 1 = 37$)
среднемесячные доходы населения составят

$$\hat{y}_{37} = y_{36} + 76,55 \cdot 1 = 3406,50 + 76,55 \cdot 1 = 3483,05 \text{ руб.}$$

(расчеты приведены в табл. 5.1).

В феврале 2002 г. ($T = 2$, $n + T = 36 + 2 = 38$) -

$$\hat{y}_{38} = y_{36} + 76,55 \cdot 2 = 3406,50 + 76,55 \cdot 2 = 3559,60 \text{ руб.}$$

В марте 2002 г. ($T = 3$, $n + T = 36 + 3 = 39$) -

$$\hat{y}_{39} = y_{36} + 76,55 \cdot 3 = 3406,50 + 76,55 \cdot 3 = 3636,15 \text{ руб.}$$

Прогнозное значение доходов населения в первом квартале 2002 г. по среднему приросту составит 10678,79 руб.

Таблица 5.1. Прогнозирование по среднему приросту

t	y_t	Δ_t	$\hat{y}_{t, модель}$	$\left \frac{y_t - \hat{y}_{t, модель}}{y_t} \right $
1	727,3		727,3	0,0000
2	868,7	141,4	803,85	0,0747
3	965,2	96,5	880,40	0,0879
4	1041,2	76	956,95	0,0809
5	954,3	-86,9	1033,49	0,0830
6	1058,5	104,2	1110,04	0,0487
7	1150,0	91,5	1186,59	0,0318
8	1072	-78	1263,14	0,1783
9	1053	-19	1339,69	0,2723
10	1539,9	486,9	1416,24	0,0803
11	1507,3	-32,6	1492,79	0,0096
12	2013	505,7	1569,33	0,2204
13	1345,3	-667,7	1645,88	0,2234
14	1587	241,7	1722,43	0,0853
15	1655,2	68,2	1798,98	0,0869
16	1723	67,8	1875,53	0,0885
17	1701,5	-21,5	1952,08	0,1473
18	1923,9	222,4	2028,63	0,0544
19	1852,6	-71,3	2105,17	0,1363
20	1987,6	135	2181,72	0,0977
21	2003,3	15,7	2258,27	0,1273
22	1971,8	-31,5	2334,82	0,1841
23	2027	55,2	2411,37	0,1896
24	2492,7	465,7	2487,92	0,0019
25	1793,8	-698,9	2564,47	0,4296
26	2146,9	353,1	2641,01	0,2302
27	2422,3	275,4	2717,56	0,1219
28	2380,8	-41,5	2794,11	0,1736
29	2315,4	-65,4	2870,66	0,2398
30	2632,8	317,4	2947,21	0,1194
31	2584,1	-48,7	3023,76	0,1701

t	y_t	Δ_t	$\hat{y}_{t, модель}$	$\left \frac{y_t - \hat{y}_{t, модель}}{y_t} \right $
32	2606,7	22,6	3100,31	0,1894
33	2547,5	-59,2	3176,85	0,2470
34	2485,6	-61,9	3253,40	0,3089
35	2611,8	126,2	3329,95	0,2750
36	3406,5	794,7	3406,50	0,0000
37			3483,05	Сумма = 5,0955
38			3559,60	
39			3636,15	

Результаты моделирования и прогнозирования по среднему приросту изображены на графике (рис. 20).

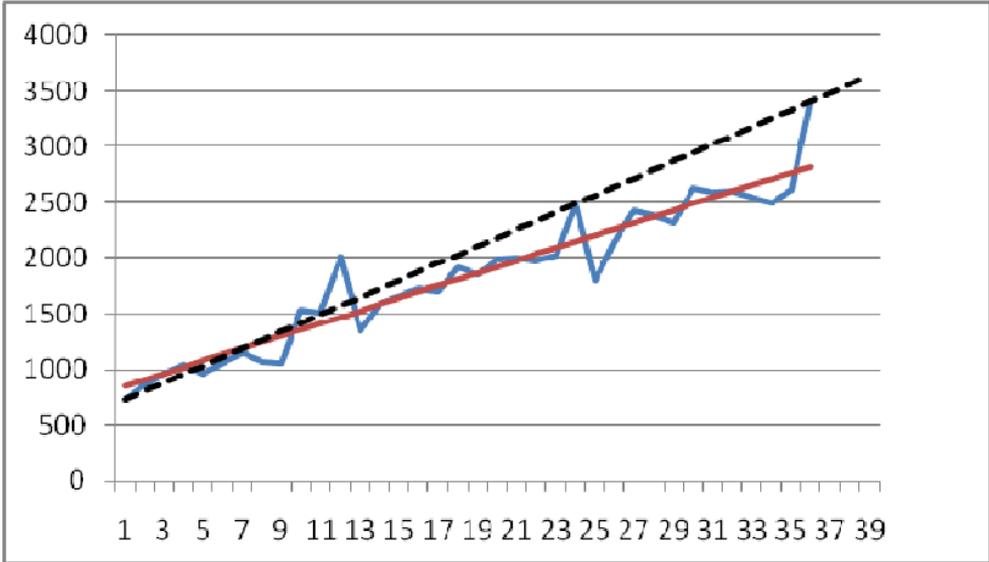


Рис. 20. Прогнозирование по среднему приросту

Как видно из графика (пунктирная линия), использование среднего прироста приводит к завышенному прогнозу доходов населения, поскольку в модели при расчете приростов учитываются циклические

колебания, имеющие наибольшую величину отклонения от тренда декабрьских и январских значений доходов.

5.2. Прогнозирование по среднему темпу роста

Прогнозирование по среднему темпу роста осуществляется, когда общая тенденция ряда динамики характеризуется показательной (экспоненциальной) кривой. В этом случае для прогнозирования определяют средний коэффициент роста \bar{P}_t , где

$$P_t = \frac{y_t}{y_{t-1}}, (t = 2, 3, \dots, n),$$

и возводят его в степень, соответствующую периоду прогнозирования T:

$$\hat{y}_{n+T} = y_n \cdot \bar{P}_t^T. \quad (5.4)$$

Для нахождения среднего коэффициента роста обычно используют среднее геометрическое:

$$\bar{P}_t^{geom} = n^{-1} \sqrt[n]{\prod_{t=2}^n P_t}. \quad (5.5)$$

Если временному ряду соответствует другая закономерность развития, то прогнозные значения, полученные по среднему темпу роста, будут отличаться от рассчитанных другими способами прогнозирования.

Точность прогноза оценивается с помощью средней относительной ошибки аппроксимации

$$\bar{A} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_{t, модель}}{y_t} \right| \cdot 100\% ,$$

где y_t - уровни ряда, $\hat{y}_{t, модель}$ - уровни ряда, найденные по модели

$$\hat{y}_{t, модель} = y_1 \cdot \bar{P}_t^{(t-1)}, (t = 2, 3, \dots, n). \quad (5.6)$$

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. найти прогноз доходов населения в первом квартале 2002 г. по среднему темпу роста. Оценить точность прогноза с помощью модельной ошибки. Результаты анализа изобразить графически

Найдем прогноз доходов населения в первом квартале 2002 г. по среднему темпу роста. Все необходимые расчеты приведены в табл. 5.2.

Темпы роста найдем по формуле

$$P_t = \frac{y_t}{y_{t-1}}, (t = 2, 3, \dots, n).$$

Средний темп роста (5.5) равен

$$\bar{P}_t^{геом} = n^{-1} \sqrt[n]{\prod_{t=2}^n P_t} = 1,05 .$$

Для его расчета использована встроенная функция Excel «СРГЕОМ».

Модельная ошибка равна

$$\bar{A} = \frac{1}{36} 0,1019 \cdot 100\% = 13,41\% .$$

Прогнозная формула (5.4):

$$\hat{y}_{n+T} = y_n \cdot 1,05^T .$$

В январе 2002 г. ($T = 1$, $n + T = 36 + 1 = 37$)
среднемесячные доходы населения составят

$$\hat{y}_{37} = y_{36} \cdot 1,05^1 = 3406,50 \cdot 1,05 = 3560,15 \text{ руб.}$$

В феврале 2002 г. ($T = 2$, $n + T = 36 + 2 = 38$) –

$$\hat{y}_{38} = y_{36} \cdot 1,05^2 = 3406,50 \cdot 1,05^2 = 3720,73 \text{ руб.}$$

В марте 2002 г. ($T = 3$, $n + T = 36 + 3 = 39$) –

$$\hat{y}_{39} = y_{36} \cdot 1,05^3 = 3406,50 \cdot 1,05^3 = 3888,55 \text{ руб.}$$

Прогнозное значение доходов населения в первом квартале 2002 г. по среднему темпу роста составит 11169,43 руб.

Таблица 5.2. Прогнозирование по среднему темпу роста

t	y_t	P_t	$\hat{y}_{t, модель}$	$\left \frac{y_t - \hat{y}_{t, модель}}{y_t} \right $
1	727,3		727,30	0,0000
2	868,7	1,19	760,10	0,1250
3	965,2	1,11	794,39	0,1770
4	1041,2	1,08	830,22	0,2026
5	954,3	0,92	867,67	0,0908
6	1058,5	1,11	906,80	0,1433
7	1150,0	1,09	947,70	0,1759
8	1072	0,93	990,45	0,0761
9	1053	0,98	1035,12	0,0170
10	1539,9	1,46	1081,81	0,2975
11	1507,3	0,98	1130,61	0,2499
12	2013	1,34	1181,60	0,4130
13	1345,3	0,67	1234,90	0,0821
14	1587	1,18	1290,60	0,1868
15	1655,2	1,04	1348,81	0,1851
16	1723	1,04	1409,65	0,1819
17	1701,5	0,99	1473,23	0,1342
18	1923,9	1,13	1539,68	0,1997
19	1852,6	0,96	1609,13	0,1314
20	1987,6	1,07	1681,71	0,1539
21	2003,3	1,01	1757,56	0,1227
22	1971,8	0,98	1836,84	0,0684
23	2027	1,03	1919,69	0,0529
24	2492,7	1,23	2006,27	0,1951
25	1793,8	0,72	2096,77	0,1689
26	2146,9	1,20	2191,34	0,0207
27	2422,3	1,13	2290,18	0,0545
28	2380,8	0,98	2393,48	0,0053
29	2315,4	0,97	2501,44	0,0803
30	2632,8	1,14	2614,26	0,0070
31	2584,1	0,98	2732,18	0,0573

t	y_t	P_t	$\hat{y}_{t, модель}$	$\left \frac{y_t - \hat{y}_{t, модель}}{y_t} \right $
32	2606,7	1,01	2855,41	0,0954
33	2547,5	0,98	2984,21	0,1714
34	2485,6	0,98	3118,81	0,2548
35	2611,8	1,05	3259,48	0,2480
36	3406,5	1,30	3406,50	0,0000
37			3560,15	Сумма = 4,8260
38			3720,73	
39			3888,55	

Результаты моделирования и прогнозирования по среднему темпу роста изображены на рис. 21.

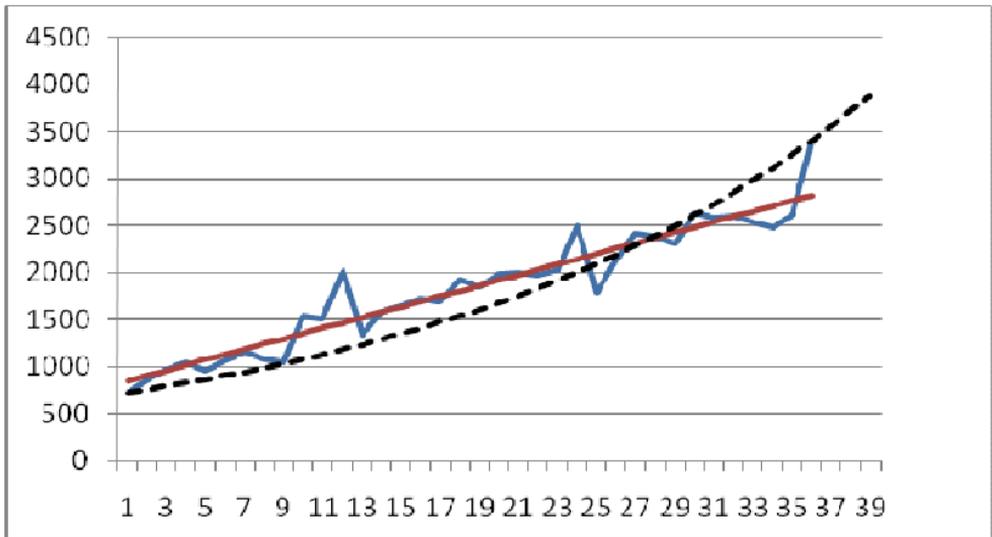


Рис. 21. Прогнозирование по среднему темпу роста

Как видно из графика (пунктирная линия), использование среднего темпа роста приводит к завышенному прогнозу доходов населения, поскольку в модели при расчете среднего темпа роста учитываются циклические колебания, имеющие наибольшую величину отклонения от тренда декабрьских и январских значений доходов.

5.3. Прогнозирование методом численного сглаживания

Хорошие результаты при краткосрочном прогнозировании значений временного ряда на один шаг вперед по времени (для $T = 1$) дает применение численного сглаживания, например, метода скользящих средних по m последним уровням ряда. Для $m = 5$ можно воспользоваться формулой

$$\hat{y}_{n+1} = \frac{1}{10} (8y_n + 5y_{n-1} + 2y_{n-2} - y_{n-3} - 4y_{n-4}). \quad (5.7)$$

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. найти прогноз доходов населения в первом квартале 2002 г. с помощью численного сглаживания. Оценить точность прогноза с помощью модельной ошибки. Результаты анализа изобразить графически

Найдем прогнозное значение доходов населения в первом квартале 2002 г. методом численного сглаживания. Воспользуемся формулой (5.7). Все необходимые расчеты приведены в табл. 5.3.

В январе 2002 г. ($n = 36$) среднемесячные доходы населения составят

$$\hat{y}_{37} = \frac{1}{10} (8y_{36} + 5y_{35} + 2y_{34} - y_{33} - 4y_{32}) = 3230,79 \text{ руб.}$$

В феврале 2002 г. ($n = 37$)

$$\hat{y}_{38} = \frac{1}{10} (8y_{37} + 5y_{36} + 2y_{35} - y_{34} - 4y_{33}) = 3542,68 \text{ руб.}$$

В марте 2002 г. (n = 38)

$$\hat{y}_{39} = \frac{1}{10} (8y_{38} + 5y_{37} + 2y_{36} - y_{35} - 4y_{34}) = 3875,42 \text{ руб.}$$

Прогнозное значение доходов населения в первом квартале 2002 г, полученное методом численного сглаживания, составит 10648,89 руб.

Модельная ошибка равна

$$\bar{A} = \frac{1}{32} 1,6756 \cdot 100\% = 5,71\% .$$

Модельные значения уровней ряда в табл. 5.3 рассчитывались по формуле

$$\hat{y}_{t, модель} = \frac{1}{10} (8y_t + 5y_{t-1} + 2y_{t-2} - y_{t-3} - 4y_{t-4}),$$

где $t = 5, 6, \dots, 36$.

Таблица 5.3. Прогнозирование методом численного сглаживания

t	y_t	$\hat{y}_{t, модель}$	$\left \frac{y_t - \hat{y}_{t, модель}}{y_t} \right $
1	727,3		
2	868,7		
3	965,2		
4	1041,2		
5	954,3	1099,29	0,1519
6	1058,5	1088,19	0,0280
7	1150	1149,91	0,0001

t	y_t	$\hat{y}_{t, модель}$	$\left \frac{y_t - \hat{y}_{t, модель}}{y_t} \right $
8	1072	1132,39	0,0563
9	1053	1120,83	0,0644
10	1539,9	1434,42	0,0685
11	1507,3	1619,19	0,0742
12	2013	2137,93	0,0621
13	1345,3	1809,01	0,3447
14	1587	1578,16	0,0056
15	1655,2	1582,5	0,0439
16	1723	1583,67	0,0809
17	1701,5	1856,92	0,0913
18	1923,9	1934,15	0,0053
19	1852,6	1949,95	0,0525
20	1987,6	2041,81	0,0273
21	2003,3	2093,97	0,0453
22	1971,8	2021,79	0,0254
23	2027	2068,36	0,0204
24	2492,7	2406,65	0,0345
25	1793,8	2088,29	0,1642
26	2146,9	2121,54	0,0118
27	2422,3	2309,98	0,0464
28	2380,8	2368,71	0,0051
29	2315,4	2594,97	0,1207
30	2632,8	2639,11	0,0024
31	2584,1	2639,76	0,0215
32	2606,7	2720,11	0,0435
33	2547,5	2668,73	0,0476
34	2485,6	2472,04	0,0055
35	2611,8	2547,43	0,0246
36	3406,5	3230,79	0,0516
37		3230,79	Сумма = 1,8276
38		3542,68	
39		3875,42	

Результаты моделирования и прогнозирования методом численного сглаживания изображены на рис. 22.

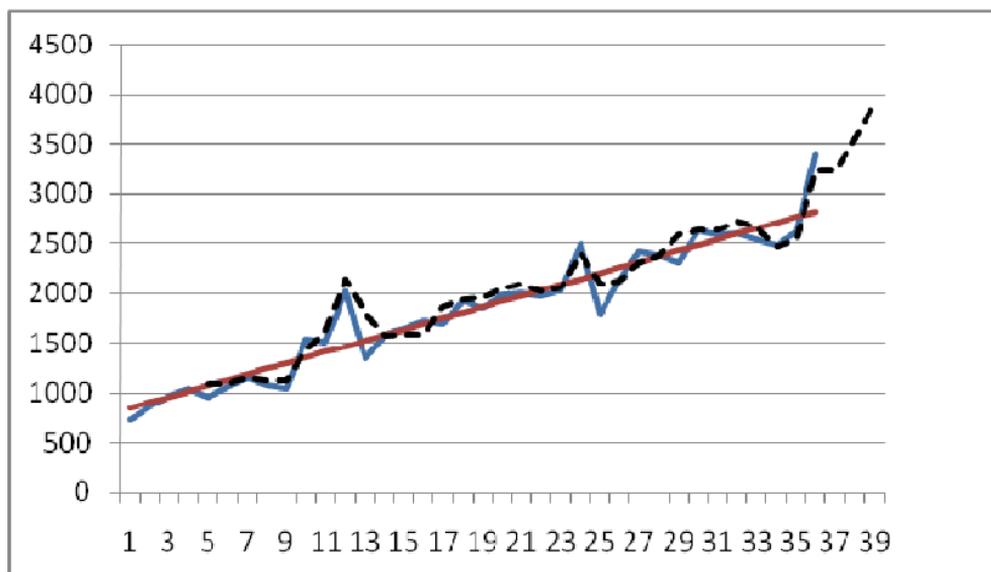


Рис. 22. Прогнозирование по методу численного сглаживания

Как видно из графика (пунктирная линия), использование метода численного сглаживания приводит к завышенному прогнозу доходов населения, поскольку в при расчете прогнозных значений уровней ряда учитывается последнее циклическое колебание, имеющее наибольшую величину отклонения от тренда декабрьского значения доходов.

Рассмотренные выше три способа прогнозирования являются самыми простейшими и приближенными.

5.4. Прогнозирование по уравнению тренда

Этот подход прогнозирования предполагает, что уровень временного ряда формируется под воздействием множества факторов, но при этом отдельно влияние каждого из них не выделяется. Следовательно, тенденция развития связана не с каким-либо фактором, а с течением времени.

На практике наибольшее распространение получил метод прогнозирования на основе *аналитического выравнивания* временного ряда (на основе уравнения тренда). При этом для получения прогнозного значения продолжают значения независимой переменной времени t за границы исследуемого периода.

Эта процедура называется экстраполяцией, т.е. продление на будущее тенденции, наблюдавшейся в прошлом. На рис. 23 прогнозная часть графика ряда изображена пунктиром.

Путем подстановки срока прогноза T (горизонта прогноза, периода упреждения) в уравнение тренда получают точечную оценку прогноза \hat{y}_{n+T} .

Например, в случае линейного тренда (3.1)

$$\hat{y}_n = b_0 + b_1 \cdot t$$

прогнозное значение уровня ряда равно

$$\hat{y}_{n+T} = b_0 + b_1 \cdot (n + T), \quad (5.8)$$

здесь n – длина ряда, T – горизонт прогноза, $n + T$ – точка прогноза.

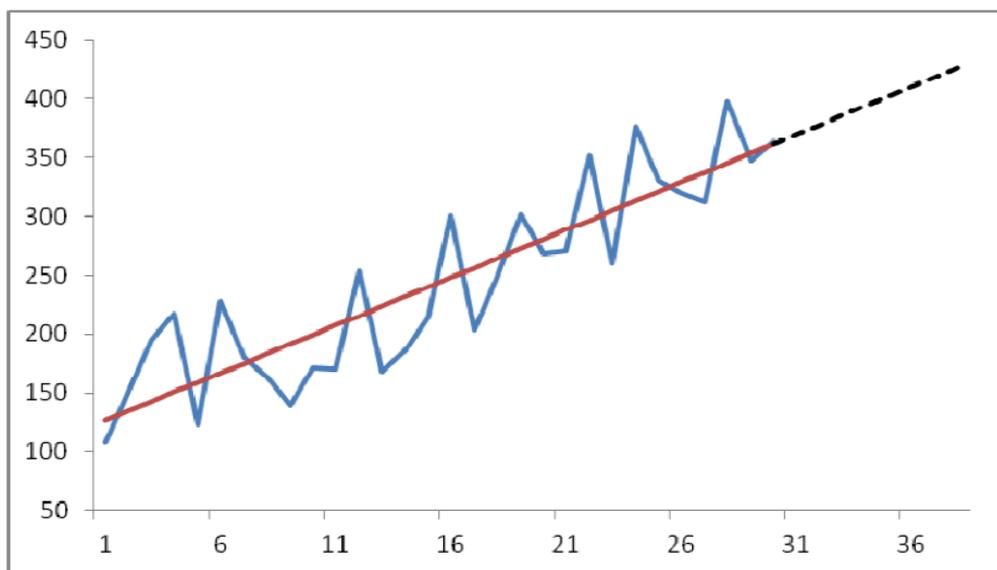


Рис. 23. Прогнозирование по уравнению тренда

Прогнозирование с использованием линейного тренда называется линейной экстраполяцией.

Прогнозные оценки, полученные подобным образом, называют точечными оценками, т.к. для каждого момента времени в период упреждения получается одно единственное значение показателя.

Поскольку экономические процессы в большинстве своем являются стохастическими, то вероятность того, что в заданный момент времени показатель будет иметь значение, определенное точечным прогнозом практически равна нулю. Т.е. полное совпадение фактического предсказываемого значения y_{n+T} и прогнозной оценки

\hat{y}_{n+T} маловероятно.

Возникновение отклонений фактических уровней временного ряда от вычисленных по уравнению тренда связано со следующими причинами:

- 1) всегда существует тренд, который дает более точные результаты при описании тенденции, по сравнению с выбранным;
- 2) тренд, выбранный для прогнозирования, содержит случайную компоненту, так как каждый уровень исходных данных обладает случайной компонентой;
- 3) выявленная тенденция характеризует движение среднего уровня временного ряда, следовательно, возможны отклонения от него.

В дополнение к точечному прогнозу определяют границы возможного изменения прогнозируемого показателя. Их еще называют доверительными интервалами или интервальными оценками прогноза. Ширина доверительных интервалов зависит от степени колеблемости исследуемого процесса, от периода упреждения, от количества наблюдений в исходном временном ряду и других факторов.

В случае линейного тренда доверительный интервал имеет вид

$$\left(\hat{y}_{n+T} - t_{1-\alpha} \cdot S_{\hat{y}} ; \hat{y}_{n+T} + t_{1-\alpha} \cdot S_{\hat{y}} \right) \quad (5.9)$$

и с вероятностью $\gamma = 1 - \alpha$ покрывает предсказываемое значение y_{n+T} . Здесь $t_{1-\alpha}$ – критическое значение распределения Стьюдента, найденное по уровню значимости α и числу степеней свободы $df = n - m - 1$,

n - длина ряда, m - число параметров уравнения тренда, T - период упреждения.

Остаточная средняя квадратическая ошибка прогноза находится по формуле

$$S_{\hat{y}} = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_{n+T} - y_t)^2}{n - m - 1}} . \quad (5.10)$$

Точность прогноза можно оценить с помощью относительной ошибки прогноза

$$\delta_{n+T} = \left| \frac{\hat{y}_{n+T} - y_{n+T}}{y_{n+T}} \right| \cdot 100\% , \quad (5.11)$$

которая может быть определена только при достижении точки прогноза (момента времени $t = n + T$), и обычно используется на этапе разработки методики прогнозирования.

Часто в качестве оценки точности прогноза берут модельную ошибку - среднюю относительную ошибку аппроксимации

$$\bar{A} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_{t, \text{модель}}}{y_t} \right| \cdot 100\% ,$$

где y_t - уровни ряда, $\hat{y}_{t, \text{модель}}$ - уровни ряда, найденные по модели.

$$\hat{y}_{t, \text{модель}} = b_0 + b_1 \cdot t, \quad (t = 1, 2, \dots, n) .$$

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. найти прогноз доходов населения в первом квартале 2002 г. с помощью уравнения тренда. Оценить точность прогноза с помощью доверительного интервала и модельной ошибки. Результаты анализа изобразить графически

Найдем прогнозные значения доходов, используя уравнение тренда $\hat{y}_t = 758,99 + 58,31 \cdot t$.

Что бы найти прогнозное значение доходов в январе 2002 г. ($T = 1$) нужно подставить точку прогноза $t = n + T = 36 + 1 = 37$ в уравнение тренда

$$\hat{y}_{37} = 758,99 + 58,31 \cdot 37 = 2916,32 .$$

Для февраля 2002 г. ($T = 2$) $t = n + T = 36 + 2 = 38$,

$$\hat{y}_{38} = 758,99 + 58,31 \cdot 38 = 2974,62 .$$

Для марта 2002 г. ($T = 3$) $t = n + T = 36 + 3 = 39$,

$$\hat{y}_{39} = 758,99 + 58,31 \cdot 39 = 3032,93 .$$

Построим интервальные оценки прогноза (доверительные интервалы) в виде (5.9)

$$\left(\hat{y}_{n+T} - t_{1-\alpha} \cdot S_{\hat{y}} ; \hat{y}_{n+T} + t_{1-\alpha} \cdot S_{\hat{y}} \right) ,$$

где $t_{1-\alpha} = 0,063$ – критическое значение распределения Стьюдента с числом степеней свободы $df = n - m - 1 = 36 - 1 - 1 = 34$, $m = 1$ – число параметров уравнения тренда при независимой переменной t , $n = 36$ – длина ряда. Расчет остаточных среднеквадратических ошибок прогноза по формуле

$$S_{\hat{y}} = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_{n+T} - y_t)^2}{n - m - 1}}$$

приведен в табл. 5.4.

Таблица 5.4. Расчет интервальных оценок прогноза

t	y_t	\hat{y}_t	$(\hat{y}_{37} - y_t)^2$	$(\hat{y}_{38} - y_t)^2$	$(\hat{y}_{39} - y_t)^2$
1	727,3	817,3	4791801	5050467	5315933
2	868,7	875,6	4192741	4434918	4683895
3	965,2	933,9	3806862	4037787	4275510
4	1041,2	992,2	3516068	3738130	3966991
5	954,3	1050,5	3849516	4081711	4320706
6	1058,5	1108,8	3451489	3671533	3898377
7	1150,0	1167,1	3119880	3329255	3545428
8	1072,0	1225,4	3401510	3619980	3845249
9	1053,0	1283,7	3471955	3692641	3920126
10	1539,9	1342,0	1894527	2058434	2229141
11	1507,3	1400,4	1985332	2153041	2327549
12	2013,0	1458,7	815983,9	924721,7	1040259
13	1345,3	1517,0	2468098	2654698	2848098
14	1587,0	1575,3	1767087	1925502	2090716
15	1655,2	1633,6	1590419	1740881	1898142
16	1723,0	1691,9	1424008	1566564	1715919
17	1701,5	1750,2	1475783	1620846	1772708

t	y_t	\hat{y}_t	$(\hat{y}_{37} - y_t)^2$	$(\hat{y}_{38} - y_t)^2$	$(\hat{y}_{39} - y_t)^2$
18	1923,9	1808,5	984894	1104022	1229949
19	1852,6	1866,8	1131497	1258939	1393181
20	1987,6	1925,1	862517,6	974217,4	1092716
21	2003,3	1983,4	833602,3	943471,3	1060139
22	1971,8	2041,7	892114,7	1005657	1125998
23	2027,0	2100,0	790887	897992,2	1011897
24	2492,7	2158,3	179452,4	232251,2	291849,3
25	1793,8	2216,6	1260047	1394346	1535445
26	2146,9	2274,9	592004,4	685127,8	785050,5
27	2422,3	2333,3	244054	305062,3	372869,9
28	2380,8	2391,6	286779,8	352627,5	425274,5
29	2315,4	2449,9	361102,7	434576,9	514850,4
30	2632,8	2508,2	80382,6	116844	160104,6
31	2584,1	2566,5	110369	152509,4	201449
32	2606,7	2624,8	95863,46	135368,4	181672,6
33	2547,5	2683,1	136026,9	182435,3	235643
34	2485,6	2741,4	185518,2	239145	299570,9
35	2611,8	2799,7	92731,37	131641,6	177351,1
36	3406,5	2858,0	240278,1	186516,5	139554
Сумма	66155,5	66155,5	56383184	61033861	65929311

Для января 2002 г.

$$S_{\hat{y}} = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_{n+T} - y_t)^2}{n - m - 1}} = \sqrt{\frac{\sum_{t=1}^{36} (\hat{y}_{37} - y_t)^2}{36 - 1 - 1}} =$$

$$= \sqrt{\frac{56383184,013}{34}} = 1287,761$$

Интервальная оценка прогноза (доверительный интервал) имеет вид

$$\begin{aligned} & (2916,318 - 0,063 \cdot 1287,761 ; 2916,318 + 0,063 \cdot 1287,761) \\ \text{или} & (2834,97 ; 2997,67) \end{aligned}$$

Для февраля 2002 г.

$$\begin{aligned} S_{\hat{y}} &= \sqrt{\frac{\sum_{t=1}^n (\hat{y}_{n+T} - y_t)^2}{n-m-1}} = \sqrt{\frac{\sum_{t=1}^{36} (\hat{y}_{38} - y_t)^2}{36-1-1}} = \\ &= \sqrt{\frac{61033861,233}{34}} = 1339,818 \end{aligned}$$

Доверительный интервал имеет вид

$$\begin{aligned} & (2974,624 - 0,063 \cdot 1339,818 ; 2974,624 + 0,063 \cdot 1339,818) \\ \text{или} & (2889,99 ; 3059,26) \end{aligned}$$

Для марта 2002 г.

$$\begin{aligned} S_{\hat{y}} &= \sqrt{\frac{\sum_{t=1}^n (\hat{y}_{n+T} - y_t)^2}{n-m-1}} = \sqrt{\frac{\sum_{t=1}^{36} (\hat{y}_{38} - y_t)^2}{36-1-1}} = \\ &= \sqrt{\frac{65929310,938}{34}} = 1392,515 \end{aligned}$$

Доверительный интервал имеет вид

$$\left(3032,931 - 0,063 \cdot 1392,515 ; 3032,931 + 0,063 \cdot 1392,515 \right)$$

или $\left(2944,96 ; 3120,90 \right)$

Таким образом, среднедушевые денежные доходы населения в январе 2002 г. Составят около 2916,32 руб. и с вероятностью 0,95 могут варьироваться от 2834,97 до 2997,67 руб.

Среднедушевые денежные доходы населения в феврале 2002 г. составят 2974,62 руб. и с вероятностью 0,95 могут принимать значения от 2889,99 до 3059,26 руб.

Среднедушевые денежные доходы населения в марте 2002 г. составят 3032,93 руб. и с вероятностью 0,95 могут принимать значения от 2944,96 руб. до 3120,90 руб.

Модельная ошибка

$$\bar{A} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \cdot 100\% = \frac{1}{36} \cdot 2,566 \cdot 100\% = 7,13\%$$

найдена на стр. 51.

Иллюстрация прогнозных значений и интервальных оценок представлена на рис. 24.

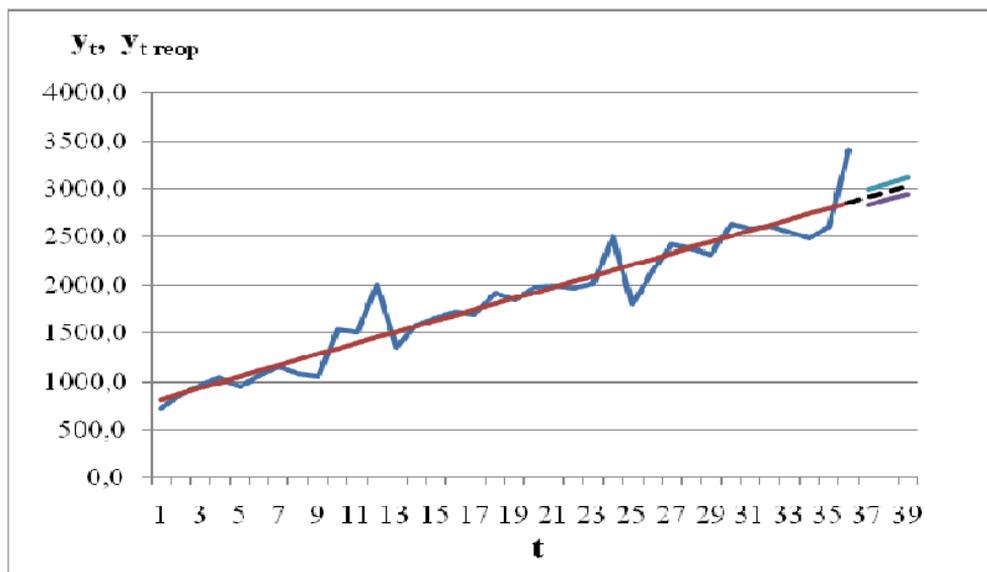


Рис. 24. Прогнозирование по линейному тренду, интервальная оценка точности прогноза

5.5. Прогнозирование по аддитивной модели с учетом сезонности

Прогнозирование по аддитивной модели с сезонной составляющей аналогично прогнозированию по тренду.

Пусть построена аддитивная модель временного ряда с сезонностью

$$y_t = \hat{y}_t + s_t + \varepsilon_t ,$$

где общая тенденция описывается линейным трендом

$$\hat{y}_t = b_0 + b_1 \cdot t , (t = 1, 2, \dots, n),$$

и найдены скорректированные сезонные составляющие внутри цикла $S_t^{сkopp}$.

Тогда прогнозное значение уровня ряда для горизонта прогноза T составит

$$\hat{y}_{n+T} = b_0 + b_1 \cdot (n + T) + S_{n+T}^{сkopp},$$

где n – длина ряда, $n + T$ – точка прогноза, $S_{n+T}^{сkopp}$ – скорректированная сезонная составляющая, соответствующая точке прогноза.

Точность прогноза можно оценить через модельную ошибку по формуле

$$\bar{A} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_{t, модель}}{y_t} \right| \cdot 100\% ,$$

где y_t – уровни ряда, $\hat{y}_{t, модель}$ – уровни ряда, найденные по модели с сезонностью

$$\hat{y}_{t, модель} = b_0 + b_1 \cdot t + S_t^{сkopp}, \quad (t = 1, 2, \dots, n). \quad (5.12)$$

Интервальная оценка прогноза по формуле (5.9) в данном случае не строится.

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. найти прогноз доходов населения в первом квартале 2002 г. по аддитивной модели с сезонной составляющей. Оценить точность прогноза с помощью модельной ошибки. Результаты анализа изобразить графически

В примере на стр. 75-83 построена аддитивная модель с сезонной составляющей

$$y_t = \hat{y}_t + s_t + \varepsilon_t$$

и линейным трендом

$$\hat{y}_t = 793,61 + 56,43 \cdot t, \quad (t = 1, 2, \dots, 36).$$

Скорректированные сезонные составляющие $S_t^{с\text{корр}}$ записаны в табл. 5.5.

Воспользовавшись формулой (5.12) получим прогноз доходов населения в первом квартале 2002 г.

Для получения прогноза среднемесячных доходов населения в январе 2002 г. ($T = 1$) подставим точку прогноза $t = n + T = 36 + 1 = 37$ в уравнение тренда и добавим скорректированную сезонную составляющую за январь

$$\hat{y}_{37} = 793,61 + 56,43 \cdot 37 + (-236,44) = 2645,25 \text{ руб.}$$

В феврале 2002 г. ($T=2$) $t = n + T = 36 + 2 = 38$, среднемесячные доходы населения по прогнозу составят $\hat{y}_{38} = 793,61 + 56,43 \cdot 38 + (-47,00) = 2891,13$ руб.

В марте 2002 г. ($T=3$) $t = n + T = 36 + 3 = 39$, $\hat{y}_{39} = 793,61 + 56,43 \cdot 39 + 108,51 = 3103,08$ руб.

Прогнозное значение доходов населения в первом квартале 2002 г. по аддитивной модели с сезонностью составит 8639,46 руб.

Модельная ошибка равна

$$\bar{A} = \frac{1}{36} 1,8824 \cdot 100\% = 5,23\% .$$

Таблица 5.5. Прогнозирование с учетом сезонности

t	y_t	$S_t^{скopp}$	$\hat{y}_{t, модель}$	$\left \frac{y_t - \hat{y}_{t, модель}}{y_t} \right $
1	727,3	-236,44	613,60	0,1563
2	868,7	-47,00	859,48	0,0106
3	965,2	108,51	1071,43	0,1101
4	1041,2	26,02	1045,37	0,0040
5	954,3	-97,52	978,27	0,0251
6	1058,5	75,58	1207,80	0,1410
7	1150,0	35,49	1224,14	0,0645
8	1072,0	-6,38	1238,71	0,1555
9	1053,0	-62,61	1238,91	0,1766
10	1539,9	-86,48	1271,48	0,1743
11	1507,3	-99,61	1314,78	0,1277
12	2013,0	390,42	1861,25	0,0754
13	1345,3	-236,44	1290,82	0,0405
14	1587,0	-47,00	1536,70	0,0317

t	y_t	$S_t^{скоп}$	$\hat{y}_{t, модель}$	$\frac{ y_t - \hat{y}_{t, модель} }{y_t}$
15	1655,2	108,51	1748,64	0,0565
16	1723,0	26,02	1722,59	0,0002
17	1701,5	-97,52	1655,49	0,0270
18	1923,9	75,58	1885,01	0,0202
19	1852,6	35,49	1901,36	0,0263
20	1987,6	-6,38	1915,92	0,0361
21	2003,3	-62,61	1916,13	0,0435
22	1971,8	-86,48	1948,70	0,0117
23	2027,0	-99,61	1992,00	0,0173
24	2492,7	390,42	2538,47	0,0184
25	1793,8	-236,44	1968,04	0,0971
26	2146,9	-47,00	2213,92	0,0312
27	2422,3	108,51	2425,86	0,0015
28	2380,8	26,02	2399,81	0,0080
29	2315,4	-97,52	2332,70	0,0075
30	2632,8	75,58	2562,23	0,0268
31	2584,1	35,49	2578,58	0,0021
32	2606,7	-6,38	2593,14	0,0052
33	2547,5	-62,61	2593,35	0,0180
34	2485,6	-86,48	2625,92	0,0565
35	2611,8	-99,61	2669,21	0,0220
36	3406,5	390,42	3215,69	0,0560
37		-236,44	2645,25	Сумма = 1,8824
38		-47,00	2891,13	
39		108,51	3103,08	

Иллюстрация прогнозных значений представлена на рис. 25.

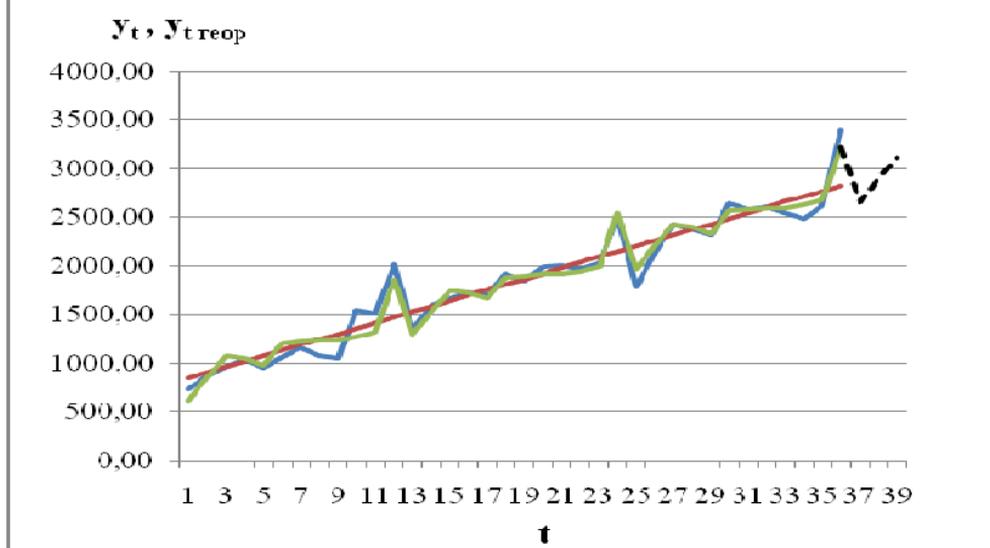


Рис. 25. Прогнозирование по аддитивной модели с сезонной составляющей и линейным трендом

Пример

Для показателя «Среднедушевые денежные доходы» (руб.) по данным с января 1999 г. по декабрь 2001 г. найден прогноз доходов населения в первом квартале 2002 г. разными методами. Сравнить результаты прогнозирования пятью способами. Для исследуемого ряда выбрать наилучшую модель прогнозирования.

Для выбора наилучшей модели прогнозирования сведем результаты, полученные в примерах в табл. 5.6.

Наилучшим методом прогнозирования для данного временного ряда является модель с учетом сезонности, поскольку она учитывает сезонные колебания и имеет наименьшую модельную ошибку – 5,23%.

Данный вывод хорошо иллюстрируется на рис. 26.

Таблица 5.6. Прогноз среднемесячных денежных доходов населения в I квартале 2002 г.

Месяц	Денежные доходы (руб.)				
	Метод прогнозирования				
	Прогнозирование по уравнению тренда	Аддитивная модель с сезонностью	Прогнозирование по среднему приросту	Прогнозирование по среднему темпу роста	Метод численного сглаживания
январь	2916,32	2645,25	3483,05	3560,15	3230,79
февраль	2974,62	2891,13	3559,60	3720,73	3542,68
март	3032,93	3103,08	3636,15	3888,55	3875,42
Итого за I квартал 2002 г. (руб.)	8923,87	8639,46	10678,79	11169,43	10648,89
Модельная ошибка (%)	7,13	5,23	14,15	13,41	5,71

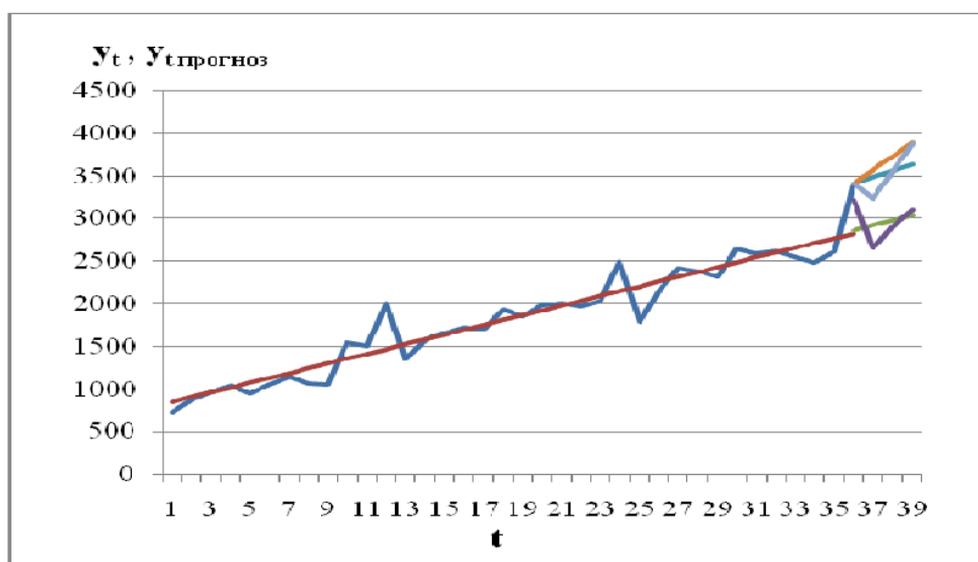


Рис. 26. Сравнение методов прогнозирования

Критические значения распределения Фишера-Снедекора при уровне значимости $\alpha = 0,05$

k1	1	2	3	4	5	6	8	12	24	∞
k2										
1	161,4	199,5	215,7	224,6	230,2	234	238,9	243,9	249	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62

k1	1	2	3	4	5	6	8	12	24	∞
k2										
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,28
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26
125	3,92	3,07	2,68	2,44	2,29	2,17	2,01	1,83	1,60	1,21
150	3,90	3,06	2,66	2,43	2,27	2,16	2,00	1,82	1,59	1,18
200	3,89	3,04	2,65	2,42	2,26	2,14	1,98	1,80	1,57	1,14
300	3,87	3,03	2,64	2,41	2,25	2,13	1,97	1,79	1,55	1,10
400	3,86	3,02	2,63	2,40	2,24	2,12	1,96	1,78	1,54	1,07
500	3,86	3,01	2,62	2,39	2,23	2,11	1,96	1,77	1,54	1,06
1000	3,85	3,00	2,61	2,38	2,22	2,10	1,95	1,76	1,53	1,03
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

Критические значения t-критерия Стьюдента

Число степеней свободы d.f.	Уровень значимости α (двусторонняя критическая область)			Число степеней свободы d.f.	Уровень значимости α (двусторонняя критическая область)		
	0,1	0,05	0,01		0,1	0,05	0,01
1	6,3138	12,706	63,657	18	1,7341	2,1009	2,8784
2	2,9200	4,3027	9,9248	19	1,7291	2,0930	2,8609
3	2,3534	3,1825	5,8409	20	1,7247	2,0860	2,8453
4	2,1318	2,7764	4,6041	21	1,7207	2,0796	2,8314
5	2,0150	2,5706	4,0321	22	1,7171	2,0739	2,8188
6	1,9432	2,4469	3,7074	23	1,7139	2,0687	2,8073
7	1,8946	2,3646	3,4995	24	1,7109	2,0639	2,7969
8	1,8595	2,3060	3,3554	25	1,7081	2,0595	2,7874
9	1,8331	2,2622	3,2498	26	1,7056	2,0555	2,7787
10	1,8125	2,2281	3,1693	27	1,7033	2,0518	2,7707
11	1,7959	2,2010	3,1058	28	1,7011	2,0484	2,7633
12	1,7823	2,1788	3,0545	29	1,6991	2,0452	2,7564
13	1,7709	2,1604	3,0123	30	1,6973	2,0423	2,7500
14	1,7613	2,1448	2,9768	40	1,6839	2,0211	2,7045
15	1,7530	2,1315	2,9467	60	1,6707	2,0003	2,6603
16	1,7459	2,1199	2,9208	120	1,6577	1,9799	2,6174
17	1,7396	2,1098	2,8982	∞	1,6449	1,9600	2,5758
Число степеней свободы d.f.	0,05	0,025	0,005	Число степеней свободы d.f.	0,05	0,025	0,005
	Уровень значимости α (односторонняя критическая область)				Уровень значимости α (односторонняя критическая область)		

**Критические значения коэффициентов
корреляции r-Пирсона**

Объем выборки n	Уровень значимости α			
	0,10	0,05	0,01	0,001
5	0,805	0,878	0,959	0,991
6	0,729	0,811	0,917	0,974
7	0,669	0,754	0,875	0,951
8	0,621	0,707	0,834	0,925
9	0,582	0,666	0,798	0,898
10	0,549	0,632	0,765	0,872
11	0,521	0,602	0,735	0,847
12	0,497	0,576	0,708	0,823
13	0,476	0,553	0,684	0,801
14	0,458	0,532	0,661	0,780
15	0,441	0,514	0,641	0,760
16	0,426	0,497	0,623	0,742
17	0,412	0,482	0,606	0,725
18	0,400	0,468	0,590	0,708
19	0,389	0,456	0,575	0,693
20	0,378	0,444	0,561	0,679
21	0,369	0,433	0,549	0,665
22	0,360	0,423	0,537	0,652
23	0,352	0,413	0,526	0,640
24	0,344	0,404	0,515	0,629
25	0,337	0,396	0,505	0,618
26	0,330	0,388	0,496	0,607
27	0,323	0,381	0,487	0,597
28	0,317	0,374	0,479	0,588
29	0,311	0,367	0,471	0,579
30	0,306	0,361	0,463	0,570
31	0,301	0,355	0,456	0,562
32	0,296	0,349	0,449	0,554
33	0,291	0,344	0,442	0,547
34	0,287	0,339	0,436	0,539
35	0,283	0,334	0,430	0,532

Объем выборки n	Уровень значимости α			
	0,10	0,05	0,01	0,001
36	0,279	0,329	0,424	0,525
37	0,275	0,325	0,418	0,519
38	0,271	0,320	0,413	0,513
39	0,267	0,316	0,408	0,507
40	0,264	0,312	0,403	0,501
41	0,260	0,308	0,398	0,495
42	0,257	0,304	0,393	0,490
43	0,254	0,301	0,389	0,484
44	0,251	0,297	0,384	0,479
45	0,248	0,294	0,380	0,474
46	0,246	0,291	0,376	0,469
47	0,243	0,288	0,372	0,465
48	0,240	0,285	0,368	0,460
49	0,238	0,282	0,365	0,456
50	0,235	0,279	0,361	0,451
51	0,233	0,276	0,358	0,447
52	0,231	0,273	0,354	0,443
53	0,228	0,271	0,351	0,439
54	0,226	0,268	0,348	0,435
55	0,224	0,266	0,345	0,432
56	0,222	0,263	0,341	0,428
57	0,220	0,261	0,339	0,424
58	0,218	0,259	0,336	0,421
59	0,216	0,256	0,333	0,418
60	0,214	0,254	0,330	0,414
61	0,213	0,252	0,327	0,411
62	0,211	0,250	0,325	0,408
63	0,209	0,248	0,322	0,405
64	0,207	0,246	0,320	0,402
65	0,206	0,244	0,317	0,399
66	0,204	0,242	0,315	0,396
67	0,203	0,240	0,313	0,393
68	0,201	0,239	0,310	0,390
69	0,200	0,237	0,308	0,388

Объем выборки n	Уровень значимости α			
	0,10	0,05	0,10	0,001
70	0,198	0,235	0,306	0,385
80	0,185	0,220	0,286	0,361
90	0,174	0,207	0,270	0,341
100	0,165	0,197	0,256	0,324
110	0,158	0,187	0,245	0,310
120	0,151	0,179	0,234	0,297
130	0,145	0,172	0,225	0,285
140	0,140	0,166	0,217	0,275
150	0,135	0,160	0,210	0,266
200	0,117	0,139	0,182	0,231
250	0,104	0,124	0,163	0,207
300	0,095	0,113	0,149	0,189
350	0,088	0,105	0,138	0,175
400	0,082	0,098	0,129	0,164
450	0,078	0,092	0,121	0,155
500	0,074	0,088	0,115	0,147
600	0,067	0,080	0,105	0,134

Критические точки распределения Дарбина-Уотсона

n	m=1		m=2		m=3		m=4		m=5	
	d _L	d _U								
6	0.61	1.40	-	-	-	-	-	-	-	-
7	0.70	1.36	0.47	1.9	-	-	-	-	-	-
8	0.76	1.33	0.56	1.78	0.37	2.29	-	-	-	-
9	0.82	1.32	0.63	1.70	0.46	2.13	-	-	-	-
10	0.88	1.32	0.70	1.64	0.53	2.02	-	-	-	-
11	0.93	1.32	0.66	1.60	0.60	1.93	-	-	-	-
12	0.97	1.33	0.81	1.58	0.66	1.86	-	-	-	-
13	1.01	1.34	0.86	1.56	0.72	1.82	-	-	-	-
14	1.05	1.35	0.91	1.55	0.77	1.78	-	-	-	-
15	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97	0,56	2,21
16	1,10	1,37	0,98	1,54	0,86	1,73	0,74	1,93	0,62	2,15
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,9	1,83	0,79	1,99
21	1,20	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,22	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,24	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,26	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,27	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,29	1,46	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,88
27	1,30	1,47	1,24	1,56	1,16	1,65	1,08	1,76	1,01	1,86
28	1,32	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,33	1,48	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,34	1,49	1,28	1,57	1,21	1,65	1,14	1,74	1,07	1,83
31	1,35	1,50	1,30	1,57	1,23	1,65	1,16	1,74	1,09	1,83
32	1,36	1,50	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82
33	1,37	1,51	1,32	1,58	1,26	1,65	1,19	1,73	1,13	1,81
34	1,38	1,51	1,33	1,58	1,27	1,65	1,21	1,73	1,15	1,81
35	1,39	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80
36	1,40	1,52	1,35	1,59	1,29	1,65	1,24	1,73	1,18	1,80
37	1,41	1,53	1,36	1,59	1,31	1,66	1,25	1,72	1,19	1,80

Продолжение прилож. 4

n	m=1		m=2		m=3		m=4		m=5	
	d _L	d _U								
38	1,42	1,54	1,37	1,59	1,32	1,66	1,26	1,72	1,21	1,79
39	1,43	1,54	1,38	1,60	1,33	1,66	1,27	1,72	1,22	1,79
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72	1,23	1,79
41	1,45	1,55	1,40	1,6	1,35	1,66	1,29	1,72	1,24	1,78
42	1,46	1,55	1,41	1,61	1,36	1,66	1,31	1,72	1,25	1,78
43	1,46	1,56	1,41	1,61	1,37	1,66	1,32	1,72	1,27	1,78
44	1,47	1,56	1,42	1,61	1,37	1,66	1,33	1,72	1,28	1,78
45	1,48	1,57	1,43	1,62	1,38	1,67	1,34	1,72	1,29	1,78
46	1,48	1,57	1,44	1,62	1,39	1,67	1,34	1,72	1,30	1,77
47	1,49	1,57	1,44	1,62	1,39	1,67	1,35	1,72	1,31	1,77
48	1,49	1,58	1,45	1,62	1,41	1,67	1,36	1,72	1,32	1,77
49	1,50	1,58	1,46	1,62	1,41	1,67	1,37	1,72	1,32	1,77
50	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72	1,34	1,77
51	1,51	1,59	1,47	1,63	1,43	1,67	1,38	1,72	1,34	1,77
52	1,51	1,59	1,47	1,63	1,43	1,68	1,39	1,72	1,35	1,77
53	1,52	1,59	1,48	1,63	1,44	1,68	1,40	1,72	1,36	1,77
54	1,52	1,60	1,48	1,64	1,45	1,68	1,41	1,72	1,37	1,77
55	1,53	1,60	1,49	1,64	1,45	1,68	1,41	1,72	1,38	1,77
56	1,53	1,60	1,49	1,64	1,46	1,68	1,42	1,72	1,38	1,77
57	1,54	1,61	1,50	1,64	1,46	1,68	1,43	1,72	1,39	1,77
58	1,54	1,61	1,50	1,65	1,47	1,68	1,43	1,72	1,39	1,77
59	1,54	1,61	1,51	1,65	1,47	1,69	1,44	1,73	1,40	1,77
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73	1,41	1,77
65	1,57	1,63	1,54	1,66	1,50	1,70	1,47	1,73	1,44	1,77
70	1,58	1,64	1,55	1,67	1,52	1,70	1,49	1,74	1,46	1,77
75	1,60	1,65	1,57	1,68	1,54	1,71	1,51	1,74	1,49	1,77
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77
85	1,62	1,67	1,60	1,70	1,57	1,72	1,55	1,75	1,52	1,77
90	1,63	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78
95	1,64	1,69	1,62	1,71	1,60	1,73	1,58	1,75	1,56	1,78
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78

Литература

1. Теория статистики: Учебник / Р.А. Шмойлова, В.Г. Минашкин, Н.А. Садовникова, Е.Б. Шувалова; Под ред. Р.А. Шмойловой. - 4-е изд., перераб. и доп. - М.: Финансы и статистика, 2006. - 656 с.

2. Эконометрика: Учебник / И.И. Елисеева, С.В. Курышева, Т.В. Костеева и др.; под ред. И.И. Елисеевой. - 2-е изд., перераб. и доп. - М.: Финансы и статистика, 2005. - 576 с.

3. Практикум по эконометрике: Учеб. пособие / И.И. Елисеева, С.В. Курышева, Н.М. Гордеенко и др.; Под ред. И.И. Елисеевой. - М.: Финансы и статистика, 2003. - 192 с.

4. Новак Эдвард. Введение в методы эконометрики. Сборник задач: Пер. с польск. / Под ред. И.И. Елисеевой. - М.: Финансы и статистика, 2004. - 248 с.

5. Балдин К.В., Быстров О.Ф., Соколов М.М. Эконометрика: Учеб. пособие для вузов - 2-е изд., перераб. и доп. - М.: ЮНИТИ-ДАНА, 2004. - 254 с.

6. Гладилин А.В., Герасимов А.Н., Громов Е.И. Эконометрика: Учеб. пособие. - М.: КНОРУС, 2006. - 232 с.

7. Афанасьев В.Н., Юзбашев М.М., Гуляева Т.И. Эконометрика: Учебник / Под ред. В.Н. Афанасьева. - М.: Финансы и статистика, 2005. - 256 с.

8. Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов. - М., Финансы и статистика, 2003.

9. Вентцель Е.С. Теория вероятностей. - М., 2002.

10. Прикладная статистика. Основы эконометрики: Учебник для вузов: В 2 т. 2-е изд. - Т.2: Айвазян С.А. Основы эконометрики. - М.: ЮНИТИ-ДАНА, 2001.

11. Боровиков В.П., Ивченко Г.И. Прогнозирование в системе STATISTICA в среде Windows. Основы теории и интенсивная практика на компьютере: Учеб. пособие. - М.: Финансы и статистика, 1999. - 384 с.

12. Берндт Э.Р. Практика эконометрики: классика и современность. - М.: ЮНИТИ-ДАНА, 2005.