

## 7. АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

### 7.1. Временной ряд и его компоненты

При рассмотрении классической модели регрессии характер экспериментальных данных, как правило, не имеет принципиального значения. Однако это оказывается не так, если условия классической модели нарушены. Методы исследования моделей, основанных на данных *пространственных выборок* и *временных рядов*, существенно отличаются. Объясняется это тем, что в отличие от пространственных выборок наблюдения во временных рядах *нельзя считать независимыми*, так как отбор данных производится не случайным образом, а в четко определенной последовательности (согласно заданной периодичности).

*Под временным рядом (динамическим рядом, или рядом динамики) подразумевается совокупность наблюдений некоторого признака  $Y$  за ряд последовательных моментов времени.*

Отдельные наблюдения называются *уровнями ряда*, которые будем обозначать  $y_t$ , при этом  $t = 1, 2, \dots, n$ , где  $n$  – число уровней или длина временного ряда.

Теоретически при исследовании экономического временного ряда  $y_t$  выделяют следующие его компоненты:

$T$  – тренд (рис. 7.1), плавно меняющаяся компонента, описывающая чистое влияние долговременных факторов, то есть длительную тенденцию изменения признака (например, плавный рост ВВП страны).

$S$  – сезонная компонента (рис. 7.2), отражающая повторяемость экономических процессов в течение не очень длительного периода (год, иногда месяц или неделя, например, объем продаж сельскохозяйственной продукции в разные времена года).

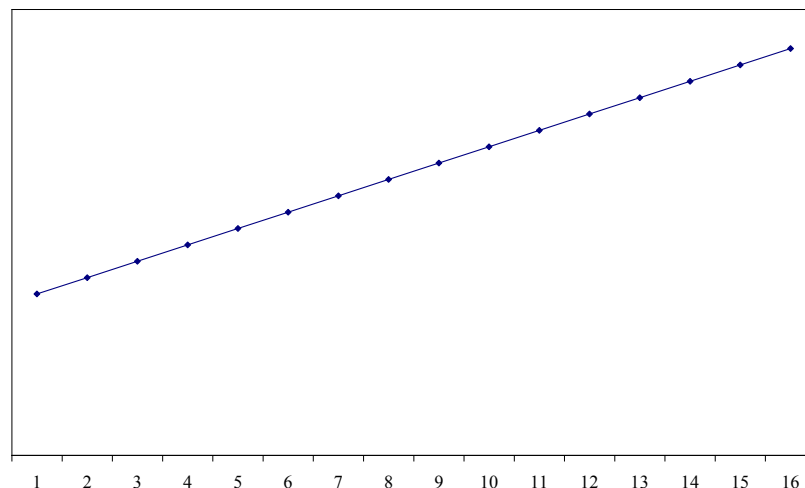


Рисунок 7.1 – Тенденция

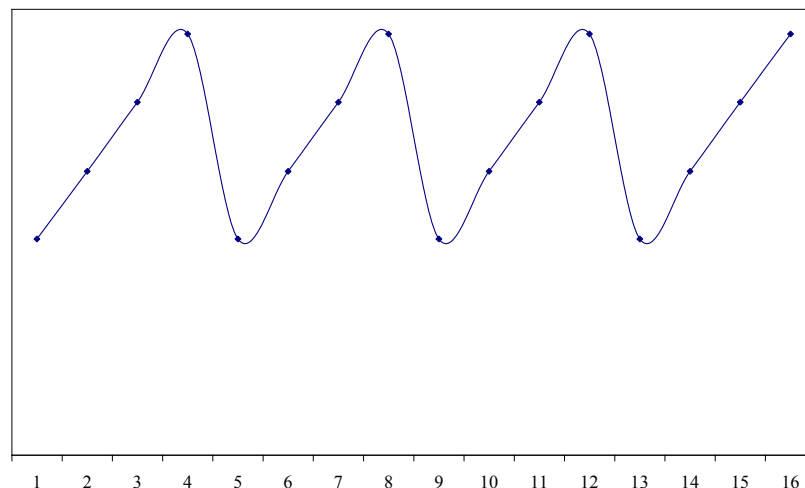


Рисунок 7.2 – Сезонные колебания

$E$  – случайная компонента (рис. 7.3), отражающая влияние не поддающихся учету и регистрации случайных факторов.

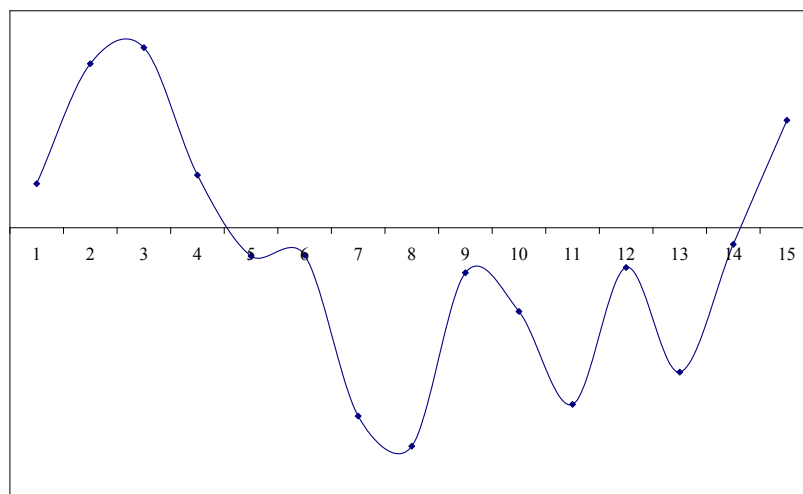


Рисунок 7.3 – Случайная компонента

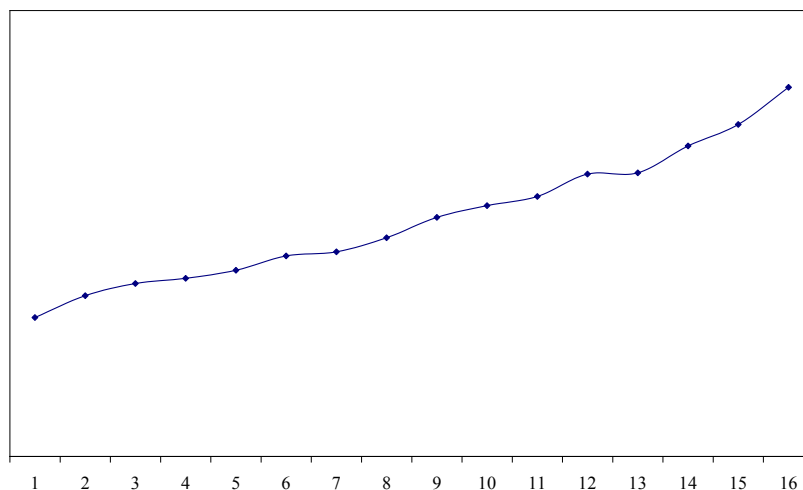


Рисунок 7.4 – Временной ряд  $y_t = f(T, E)$

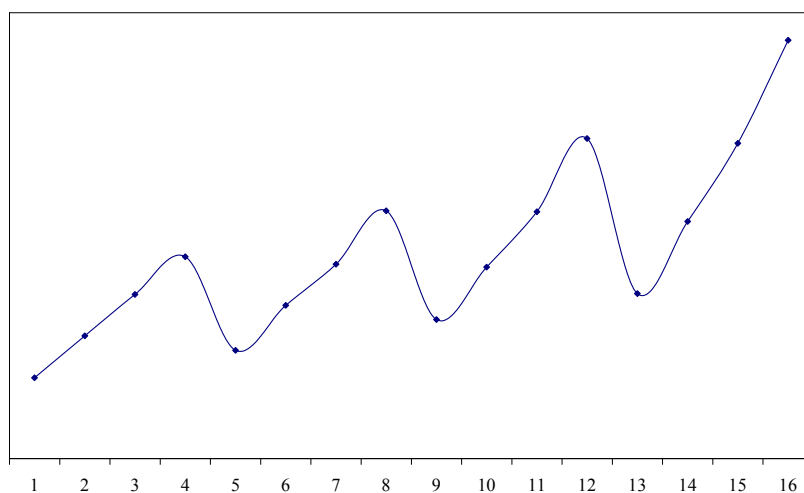


Рисунок 7.5 – Временной ряд  $y_t = f(T, S, E)$

Таким образом, любой экономический временной ряд можно представить в виде сочетания

входящих в него компонент. На рисунках 7.4 и 7.5 представлены временные ряды типа  $y_t = f(T, E)$  и типа  $y_t = f(T, S, E)$  соответственно.

## 7.2. Автокорреляция уровней ряда

Структура любого временного ряда определяется тем, насколько тесно взаимосвязаны между собой текущие наблюдения  $y_t$  и предыдущие  $y_{t-\tau}$ . Корреляционную зависимость между последовательными уровнями временного ряда называют *автокорреляцией уровней ряда*.

Количественно ее можно измерить с помощью *линейного коэффициента корреляции* между уровнями исходного временного ряда  $y_t$  и уровнями этого же ряда, сдвинутыми на несколько ( $\tau$ ) шагов во времени  $y_{t-\tau}$ .

Число моментов времени  $\tau$ , на которое сдвигается исходный временной ряд, для расчета коэффициента автокорреляции, называется *временным лагом*. То есть если требуется оценить тесноту взаимосвязи между текущими  $y_t$  и соседними предыдущими  $y_{t-1}$  наблюдениями, то лаг будет равен единице ( $\tau = 1$ ). Число периодов  $\tau$ , на которые сдвигается исходный временной ряд, определяет порядок коэффициента автокорреляции.

Таким образом, формула для расчета коэффициента автокорреляции порядка  $\tau$  будет выглядеть следующим образом:

$$r_\tau = \frac{\sum (y_t - \bar{y})(y_{t-\tau} - \bar{y})}{\sum (y_t - \bar{y})^2}.$$

То есть коэффициенты автокорреляции 1-го и 2-го порядков можно рассчитать с помощью формул

$$r_1 = \frac{\sum (y_t - \bar{y})(y_{t-1} - \bar{y})}{\sum (y_t - \bar{y})^2}.$$

$$r_2 = \frac{\sum (y_t - \bar{y})(y_{t-2} - \bar{y})}{\sum (y_t - \bar{y})^2}.$$

С увеличением лага  $\tau$  число пар значений  $n - \tau$ , по которым рассчитывается коэффициент автокорреляции, уменьшается, поэтому лаг  $\tau$  должен быть таким, чтобы число  $n - \tau$  было достаточным для определения статистически значимого  $r^\tau$ . Обычно ориентируются на соотношение  $\tau \leq n/4$ , то есть величина лага не должна превышать четверти длины исследуемого временного ряда.

## 7.3. Определение структуры временного ряда

Последовательность коэффициентов автокорреляции первого, второго и т.д. порядков называется *автокорреляционной функцией временного ряда*. График зависимости ее значений от величины лага (порядка коэффициента автокорреляции) называется *коррелограммой*.

Анализ автокорреляционной функции и коррелограммы позволяет определить лаг, при котором автокорреляция наиболее высокая, то есть лаг, при котором связь между текущими и предыдущими уровнями ряда наиболее тесная. С помощью подобных исследований автокорреляционной функции и коррелограммы можно выявить структуру временного ряда.

Если наиболее высоким оказался коэффициент автокорреляции 1-го порядка  $r^1$  и с увеличением лага значения  $r^\tau$  медленно затухают, исследуемый ряд содержит тенденцию  $T$  и случайные колебания  $E$  (ряд типа  $y_t = f(T, E)$ ).

Если коэффициент автокорреляции первого порядка строго равен единице ( $r^1 = 1$ ) и с увеличением лага значения  $r^\tau$  медленно затухают, исследуемый ряд содержит только тенденцию  $T$

(ряд типа  $y_t = f(T)$ ). Подобные временные ряды в экономике встречаются крайне редко.

Если наиболее высоким оказался коэффициент автокорреляции 4-го порядка  $r^4$  при умеренном значении  $r^1$ , временной ряд наряду с тенденцией содержит сезонные колебания  $S$  с периодичностью в четыре момента времени (ряд типа  $y_t = f(T, S, E)$ ).

Если коэффициент автокорреляции 4-го порядка строго равен единице ( $r^4 = 1$ ), исследуемый временной ряд содержит только сезонные колебания  $S$  (ряд типа  $y_t = f(S)$ ). Подобные временные ряды в экономике встречаются крайне редко.

Если наиболее высоким оказался коэффициент автокорреляции порядка  $\tau$ , временной ряд содержит сезонные колебания в  $\tau$  моментов времени.

Если ни один из коэффициентов не является значимым или автокорреляционная функция обрывается после коэффициента  $r^1$  или  $r^2$ , ряд содержит только случайные колебания  $E$  (стационарный ряд  $y_t = f(E)$ ).

*Важнейшая классическая задача при исследовании экономических временных рядов заключается в выявлении и количественной оценке основной тенденции развития  $T$ , сезонной компоненты  $S$  и случайных колебаний  $E$ .*

#### 7.4. Моделирование тенденции

Одним из наиболее распространенных способов моделирования временных рядов типа  $y_t = f(T, E)$  является построение *аналитической функции* (уравнения линейной или нелинейной регрессии. уравнения тренда), характеризующей зависимость уровней ряда  $y_t$  от времени  $t = 1, 2, \dots, n$ :

$$y_t = f(t) + e_t.$$

Этот способ моделирования называют *аналитическим выравнением временного ряда*. Выравнивание производится путем расчета теоретических (выровненных) значений  $\hat{y}_t$  моделируемого показателя  $Y$  с помощью имеющейся аналитической функции.

При выполнении аналитического выравнивания важным этапом является выбор формы уравнения тренда. Наиболее часто используется линейная функция

$$y_t = b_0 + b_1 t + e_t.$$

Оценка параметров линейного тренда производится с помощью следующей системы уравнений:

$$\begin{cases} \sum y_t = nb_0 + b_1 \sum t \\ \sum y_t t = b_0 \sum t + b_1 \sum t^2 \end{cases}.$$

Однако динамика экономических процессов может демонстрировать и нелинейный характер. В таких случаях для моделирования тенденции необходимо использовать соответствующие нелинейные функции.

Если после некоторого роста значения временного ряда начинают убывать (рис. 7.6), для его моделирования целесообразно использовать *полином второй степени*

$$y_t = b_0 + b_1 t + b_2 t^2 + e_t.$$

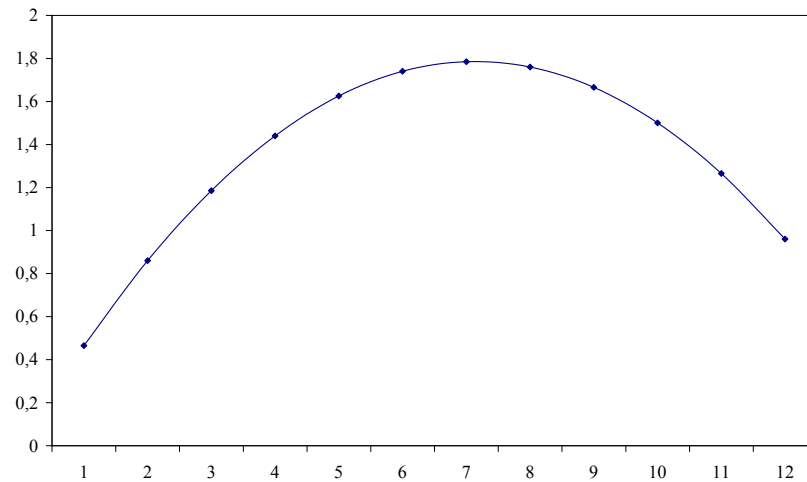


Рисунок 7.6 – Полиномиальная динамика

Для оценки параметров полиномиального тренда необходимо в уравнение линейного тренда добавить еще одну объясняющую переменную  $t^2$ . Тогда расчет параметров будет производиться с помощью следующей системы уравнений:

$$\begin{cases} \sum y_t = nb_0 + b_1 \sum t + b_2 \sum t^2 \\ \sum y_t t = b_0 \sum t + b_1 \sum t^2 + b_2 \sum t^3 \\ \sum y_t t^2 = b_0 \sum t^2 + b_1 \sum t^3 + b_2 \sum t^4 \end{cases}.$$

Если моделируемый показатель демонстрирует замедляющуюся тенденцию к снижению или росту (рис. 7.7), для ее моделирования целесообразно использовать *гиперболическую функцию*:

$$y_t = b_0 + \frac{b_1}{t} + e_t.$$

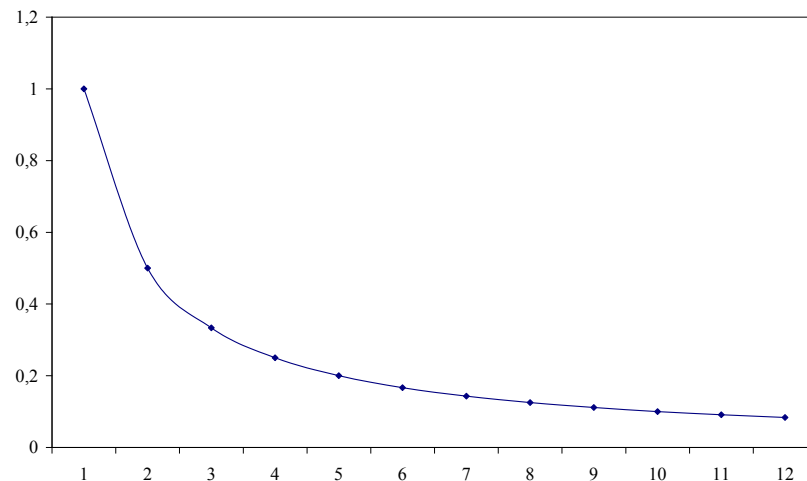


Рисунок 7.7 – Замедляющаяся тенденция к снижению

Для оценки параметров гиперболического тренда в уравнении линейного тренда вместо значений переменной  $t$  необходимо использовать преобразованные значения  $1/t$ . Тогда расчет параметров будет производиться с помощью следующей системы уравнений:

$$\begin{cases} \sum y_t = nb_0 + b_1 \sum \frac{1}{t} \\ \sum y_t \frac{1}{t} = b_0 \sum \frac{1}{t} + b_1 \sum \left(\frac{1}{t}\right)^2 \end{cases}$$

Если динамика изучаемого процесса демонстрирует замедляющийся рост (рис. 7.8), для моделирования его тенденции целесообразнее использовать *степенную функцию*

$$y_t = b_0 t^{b_1} e_t.$$

$$\ln y_t = \ln b_0 + b_1 \cdot \ln t + \ln e_t$$

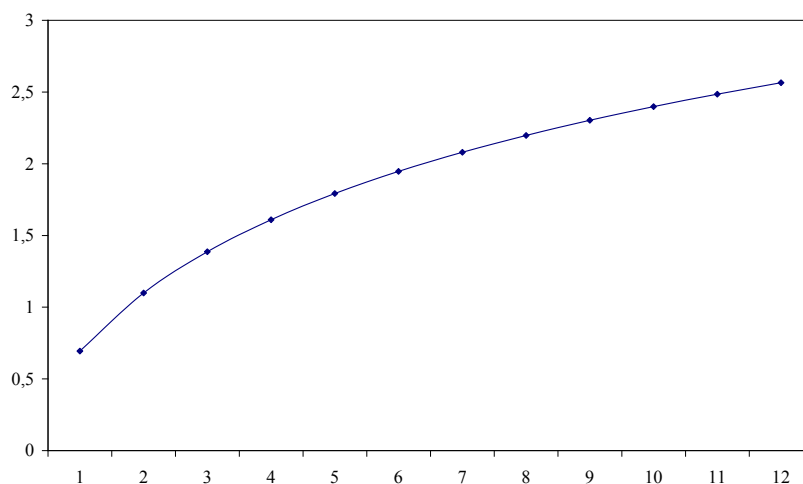


Рисунок 7.8 – Замедляющийся рост

Для оценки параметров степенного тренда необходимо в качестве зависимой переменной использовать логарифмированные значения временного ряда  $\ln y_t$ , а в качестве объясняющей переменной – логарифмированные значения времени  $\ln t$ . Тогда расчет параметров будет производиться с помощью следующей системы уравнений:

$$\begin{cases} \sum \ln y_t = n \ln b_0 + b_1 \sum \ln t \\ \sum \ln y_t \ln t = \ln b_0 \sum \ln t + b_1 \sum (\ln t)^2 \end{cases}$$

Если динамика изучаемого процесса демонстрирует ускоренный рост (рис. 7.9), для моделирования его тенденции целесообразнее использовать *показательную функцию*

$$y_t = b_0 b_1^t e_t$$

$$\ln y_t = \ln b_0 + \ln b_1 \cdot t + \ln e_t$$

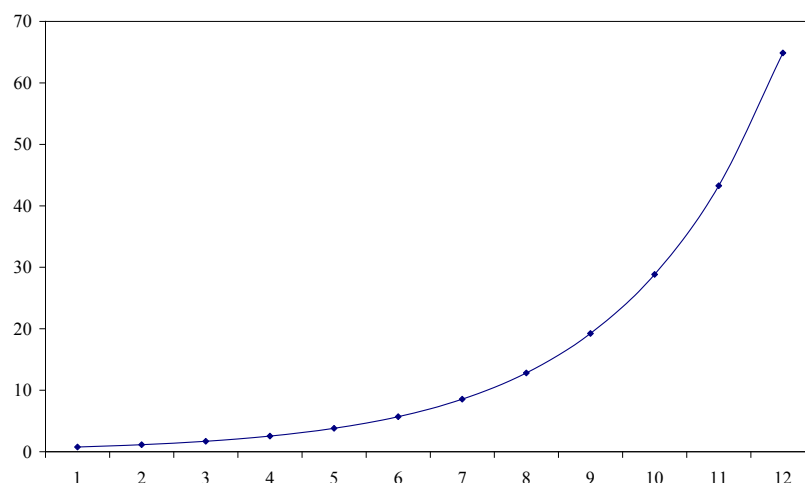


Рисунок 7.9 – Ускоренный рост

Для оценки параметров показательного тренда в качестве зависимой переменной необходимо использовать логарифмированные значения временного ряда  $\ln y_t$ , а в качестве объясняющей переменной – значения условного времени  $t = 1, 2, \dots, n$ . Тогда расчет параметров будет производиться с помощью следующей системы уравнений:

$$\begin{cases} \sum \ln y_t = n \ln b_0 + \ln b_1 \sum t \\ \sum \ln y_t \cdot t = \ln b_0 \sum t + \ln b_1 \sum t^2 \end{cases}$$

### 7.5. Выбор оптимальной формы уравнения тренда

Выбор функции  $f(t)$ , наилучшим образом описывающей моделируемый показатель, производят на основе теоретического анализа изучаемого явления (теоретические предпосылки изучаемого процесса), визуального наблюдения за динамикой процесса (на основе графического изображения временного ряда) или на основе показателей качества (показателя детерминации, средней ошибки аппроксимации) полученной трендовой модели.

Из двух функций предпочтение обычно отдается той, при которой сумма квадратов отклонений фактических данных  $y_t$  от расчетных  $\hat{y}_t$  минимальна, то есть той функции, для которой коэффициент или индекс нелинейной детерминации максимален, а средняя ошибка аппроксимации минимальна.

Однако этот принцип нельзя доводить до абсурда. Для любого ряда из  $n$  точек можно подобрать полином  $(n-1)$ -й степени, проходящий через все точки, и соответственно с минимальной (нулевой) суммой квадратов отклонений, но в этом случае, неправильно говорить о выделении основной тенденции, учитывая случайный характер этих точек. Поэтому предпочтение следует отдавать более простым функциям.

При выборе полиномиальной функции может быть полезен метод последовательных разностей, состоящий в вычислении разностей 1-го порядка  $\Delta_t = y_t - y_{t-1}$ , 2-го порядка  $\Delta_t^2 = \Delta_t - \Delta_{t-1}$  и т. д. Порядок разностей, при котором они будут примерно одинаковыми, принимается за степень полинома.

### 7.6. Прогнозирование по уравнению тренда

Одна из важнейших задач анализа временных рядов состоит в прогнозировании развития изучаемого процесса. При этом исходят из того, что тенденция развития, установленная в прошлом,

может быть распространена (экстраполирована) на будущий период.

Точечный прогноз  $\hat{y}_{n+k}$  выполняется путем подстановки  $(n+k)$ -го значения  $t_{n+k}$  объясняющей переменной  $t$  в построенное уравнение тренда. Так, например, точечный прогноз с помощью линейного тренда выполняется следующим образом:

$$y_{n+k} = b_0 + b_1 t_{n+k}.$$

Моделирование и прогнозирование стохастических процессов, как известно, предполагают наряду с точечной построение интервальной оценки прогноза. Интервальный прогноз в данном случае получают путем расчета доверительного интервала

$$\hat{y}_{n+k} - S_{y_{n+k}} \cdot t_{\alpha;df} \leq \hat{y}_{n+k} \leq \hat{y}_{n+k} + S_{y_{n+k}} \cdot t_{\alpha;df}$$

для каждого значения  $\hat{y}_{t+k}$  ( $k$  – число моментов времени, на величину которого выполняется прогноз),

где

$$S_{y_{n+k}} = S_e \sqrt{1 + \frac{1}{n} + \frac{(t_{n+k} - \bar{t})^2}{\sum (t_i - \bar{t})^2}} - \text{стандартная ошибка прогноза.}$$

$$S_e = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{n - p - 1}} - \text{стандартная ошибка регрессии.}$$

### Пример моделирования тенденции временного ряда

Имеются статистические данные о ежемесячной динамике оборота розничной торговли в Республике Бурятия за 2008 г. (табл. 7.1):

Таблица 7.1

Месяц	1	2	3	4	5	6
$Y_t$	3,74	4,32	4,65	4,79	5,01	5,40
Месяц	7	8	9	10	11	12
$Y_t$	5,51	5,88	6,43	6,75	6,99	7,60

где  $Y$  – оборот розничной торговли, млрд. руб.



### Задание:

1. Произвести анализ структуры имеющегося временного ряда визуально и с помощью автокорреляционной функции.
2. Построить трендовые модели с использованием изученных ранее функций, проверить их качество, выполнить точечные и интервальные прогнозы оборота розничной торговли на первый квартал 2009 г.
3. Выбрать тренд, наилучшим образом описывающий имеющийся временной ряд.

### Решение:

1. Анализируя график динамики оборота розничной торговли (рис. 7.10), можно предположить, что временной ряд содержит тенденцию  $T$  и случайные колебания  $E$  (временной ряд типа  $y_t = f(T, E)$ ).

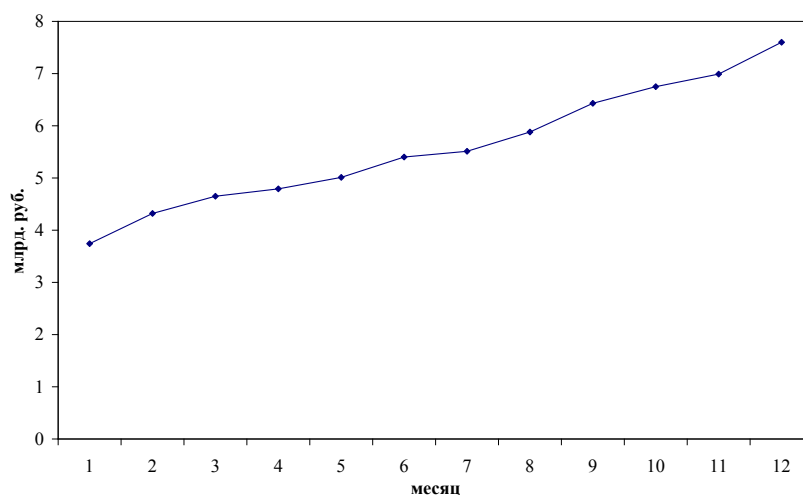


Рисунок 7.10 – Динамика оборота розничной торговли, млн.руб.

Выводы, сделанные на основе графика, требуется подтвердить количественно, с помощью коэффициентов автокорреляции 1-го, 2-го и т.д. порядков. Пользуясь данными итоговой строки таблицы 7.2, рассчитаем их значения:

$$r_1 = \frac{\sum (y_t - \bar{y})(y_{t-1} - \bar{y})}{\sum (y_t - \bar{y})^2} = \frac{10,517}{15,073} = 0,698$$

$$r_2 = \frac{\sum (y_t - \bar{y})(y_{t-2} - \bar{y})}{\sum (y_t - \bar{y})^2} = \frac{7,22}{15,073} = 0,479$$

$$r_3 = \frac{\sum (y_t - \bar{y})(y_{t-3} - \bar{y})}{\sum (y_t - \bar{y})^2} = \frac{4,133}{15,073} = 0,274$$

$$r_4 = \frac{\sum (y_t - \bar{y})(y_{t-4} - \bar{y})}{\sum (y_t - \bar{y})^2} = \frac{0,92}{15,073} = 0,061$$

Таблица 7.2

$t$	$y_t$	$(y_t - \bar{y})^2$	$y_{t-1}$	1	$y_{t-2}$	2	$y_{t-3}$	3	$y_{t-4}$	4
1	3,74	3,419	-	-	-	-	-	-	-	-
2	4,32	1,611	3,74	2,347	-	-	-	-	-	-
3	4,65	0,882	4,32	1,192	3,74	1,737	-	-	-	-
4	4,79	0,639	4,65	0,751	4,32	1,014	3,74	1,478	-	-
5	5,01	0,335	4,79	0,463	4,65	0,544	4,32	0,735	3,74	1,071
6	5,4	0,036	5,01	0,110	4,79	0,151	4,65	0,178	4,32	0,240
7	5,51	0,006	5,4	0,015	5,01	0,046	4,79	0,063	4,65	0,074
8	5,88	0,085	5,51	-0,023	5,4	-0,055	5,01	-0,168	4,79	-0,232
9	6,43	0,707	5,88	0,245	5,51	-0,067	5,4	-0,159	5,01	-0,487
10	6,75	1,348	6,43	0,976	5,88	0,338	5,51	-0,092	5,4	-0,220
11	6,99	1,962	6,75	1,626	6,43	1,178	5,88	0,407	5,51	-0,111
12	7,6	4,043	6,99	2,817	6,75	2,334	6,43	1,691	5,88	0,585
Сумма	67,07	15,073		10,517		7,220		4,133		0,920

В таблице 7.2 столбцы, озаглавленные цифрами 1, 2, 3, 4 рассчитывались следующим образом:

Столбец «1» –  $(y_t - \bar{y})(y_{t-1} - \bar{y})$ .

Столбец «2» –  $(y_t - \bar{y})(y_{t-2} - \bar{y})$ .

Столбец «3» –  $(y_t - \bar{y})(y_{t-3} - \bar{y})$ .

Столбец «4» –  $(y_t - \bar{y})(y_{t-4} - \bar{y})$ .

Как уже было отмечено, последовательность коэффициентов автокорреляции образует автокорреляционную функцию, которая графически представляется в виде коррелограммы (рис. 7.11).

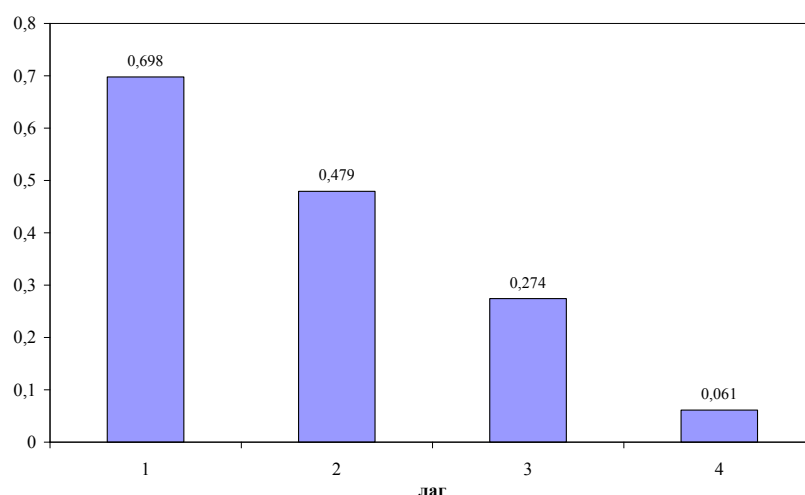


Рисунок 7.11 – Коррелограмма

Из графика коррелограммы и значений коэффициентов автокорреляции видно, что наиболее тесно связаны между собой текущие  $y_t$  и предыдущие  $y_{t-1}$  уровни временного ряда (коэффициент  $r_1 = 0,698$  является максимальным). Это свидетельствует о том, что наблюдаемый процесс включает в себя две компоненты: тенденцию  $T$  и случайные колебания  $E$ , то есть предположения, сделанные на основе графика динамики наблюдаемых значений (рис. 7.10), подтверждаются.

2. Для начала построим уравнение *линейного тренда*. Произведем необходимые расчеты для получения системы линейных уравнений (табл. 7.3). Пользуясь предпоследней строкой таблицы 7.3, получаем:

$$\begin{cases} 67,07 = 12b_0 + 78b_1 \\ 482,01 = 78b_0 + 650b_1 \end{cases},$$

тогда

$$b_0 = \frac{\Delta_{b_0}}{\Delta} = \frac{5998,72}{1716} = 3,496. \quad b_1 = \frac{\Delta_{b_1}}{\Delta} = \frac{552,66}{1716} = 0,322.$$

Уравнение линейного тренда запишется следующим образом:

$$y_t = 3,496 + 0,322 \cdot t + e_t.$$

Параметр  $b_1 = 0,322$  здесь показывает средний абсолютный прирост оборота розничной торговли за двенадцать месяцев, то есть в среднем за год ежемесячный прирост оборота розничной торговли составил 0,322 млрд. руб.

Таблица 7.3

Месяц	$y_t$	$t$	$t^2$	$y_t \cdot t$	$(t - \bar{t})^2$
1	3,74	1	1	3,74	30,25
2	4,32	2	4	8,64	20,25
3	4,65	3	9	13,95	12,25
4	4,79	4	16	19,16	6,25
5	5,01	5	25	25,05	2,25
6	5,4	6	36	32,4	0,25
7	5,51	7	49	38,57	0,25
8	5,88	8	64	47,04	2,25
9	6,43	9	81	57,87	6,25
10	6,75	10	100	67,5	12,25
11	6,99	11	121	76,89	20,25
12	7,6	12	144	91,2	30,25
<b>Сумма</b>	<b>67,07</b>	<b>78</b>	<b>650</b>	<b>482,01</b>	<b>143</b>
<b>Среднее</b>	<b>5,589</b>	<b>6,5</b>	<b>54,167</b>	<b>40,168</b>	<b>11,917</b>

Процедура выравнивания – это не что иное, как нахождение теоретических значений  $\hat{y}_t$  путем постановки значений объясняющей переменной  $t$  в имеющееся уравнение тренда (табл. 7.4). Первые два выровненных значения были получены следующим образом:

$$\hat{y}_1 = 3,496 + 0,322 \cdot 1 = 3,828.$$

$$\hat{y}_2 = 3,496 + 0,322 \cdot 2 = 4,14.$$

Остальные значения рассчитывались аналогичным образом.

Для проверки качества полученного тренда рассчитаем факторную  $Q_R$  и остаточную  $Q_e$  суммы квадратов отклонений. Пользуясь последней строкой таблицы 7.4, получаем:

$$R^2 = \frac{Q_R}{Q_R + Q_e} = \frac{14,833}{14,833 + 0,241} = 0,984.$$

$$\bar{A} = \frac{1}{n} \sum \frac{|y_t - \hat{y}_t|}{y_t} \cdot 100\% = \frac{1}{12} \cdot 25,384 = 2,115.$$

Таблица 7.4

$t$	$\hat{y}_t$	$(\hat{y}_t - \bar{y})^2$	$(y_t - \hat{y}_t)^2$	$ y_t - \hat{y}_t  / y_t \cdot 100$
1	3,818	3,138	0,00606	2,081
2	4,140	2,100	0,03244	4,169
3	4,462	1,271	0,03536	4,044
4	4,784	0,648	0,00004	0,125
5	5,106	0,233	0,00923	1,918
6	5,428	0,026	0,00079	0,521
7	5,750	0,026	0,05770	4,359
8	6,072	0,233	0,03696	3,270
9	6,394	0,648	0,00127	0,555
10	6,716	1,271	0,00113	0,498
11	7,038	2,100	0,00235	0,693
12	7,361	3,138	0,05735	3,151
<b>Сумма</b>		<b>14,833</b>	<b>0,241</b>	<b>25,384</b>

По условию задачи последнее значение переменной  $t = 12$  соответствует последнему месяцу 2008 г. Следовательно, для прогноза оборота розничной торговли на первый квартал 2009 г. необходимо продлить переменную  $t$  на три (три месяца в квартале) момента времени вперед ( $t_{n+k} = 13, 14, 15$ ) и подставить эти значения в уравнение тренда. Прогнозные значения оборота розничной торговли на первый квартал 2009 г. будут следующими:

$$\hat{y}_{13} = 3,496 + 0,322 \cdot 13 = 7,683.$$

$$\hat{y}_{14} = 3,496 + 0,322 \cdot 14 = 8,005.$$

$$\hat{y}_{15} = 3,496 + 0,322 \cdot 15 = 8,327.$$

Выполненный прогноз, как известно, является точечным. Для получения интервального прогноза рассчитаем доверительные интервалы. Здесь стандартная ошибка прогноза будет рассчитываться по формуле:

$$S_{y_{n+k}} = S_e \sqrt{1 + \frac{1}{n} + \frac{(t_{n+k} - \bar{t})^2}{\sum (t_i - \bar{t})^2}}.$$

Так как для данного случая  $t_{0,95;10} = 2,23$ , нижние границы доверительных интервалов для соответствующих прогнозных значений будут следующими:

$$\hat{y}_{13} - S_{y_{13}} \cdot t_{0,95;10} = 7,683 - 0,155 \cdot \sqrt{1 + \frac{1}{12} + \frac{(13 - 6,5)^2}{143}} \cdot 2,23 = 7,276.$$

$$\hat{y}_{14} - S_{y_{14}} \cdot t_{0,95;10} = 8,005 - 0,155 \cdot \sqrt{1 + \frac{1}{12} + \frac{(14 - 6,5)^2}{143}} \cdot 2,23 = 7,584.$$

$$\hat{y}_{15} - S_{y_{15}} \cdot t_{0,95;10} = 8,327 - 0,155 \cdot \sqrt{1 + \frac{1}{12} + \frac{(15 - 6,5)^2}{143}} \cdot 2,23 = 7,891.$$

а верхние – следующими:

$$\hat{y}_{13} + S_{y_{13}} \cdot t_{0,95;10} = 7,683 + 0,155 \cdot \sqrt{1 + \frac{1}{12} + \frac{(13 - 6,5)^2}{143}} \cdot 2,23 = 8,089.$$

$$\hat{y}_{14} + S_{y_{14}} \cdot t_{0,95;10} = 8,005 + 0,155 \cdot \sqrt{1 + \frac{1}{12} + \frac{(14 - 6,5)^2}{143}} \cdot 2,23 = 8,425.$$

$$\hat{y}_{15} + S_{y_{15}} \cdot t_{0,95;10} = 8,327 + 0,155 \cdot \sqrt{1 + \frac{1}{12} + \frac{(15 - 6,5)^2}{143}} \cdot 2,23 = 8,763.$$

Полученные верхние границы доверительного интервала можно классифицировать как *оптимистический* прогноз, нижние – как *пессимистический*, а точечный прогноз – как *наиболее вероятный*. Представим полученные результаты расчетов на графике (рис. 7.12).

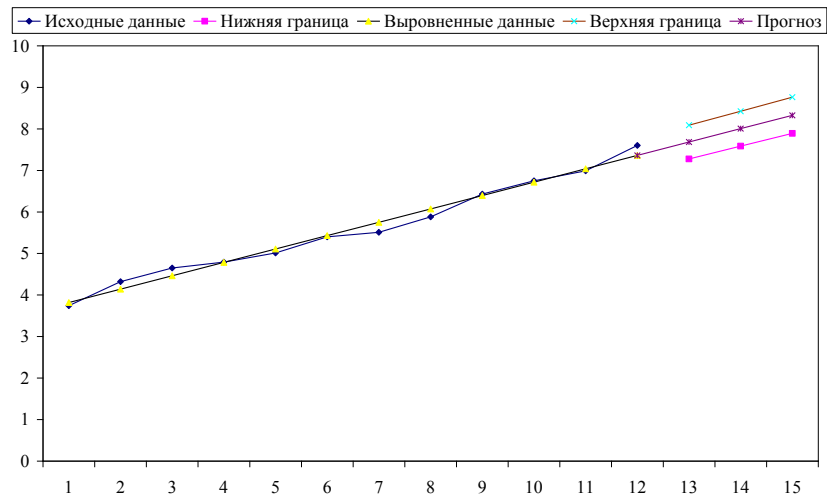


Рисунок 7.12 – Соотношение наблюдаемых  $y_t$ , выровненных  $\hat{y}_t$  и прогнозных  $\hat{y}_{n+k}$  значений для линейной функции

Теперь протестируем, насколько качественно аппроксимируют имеющиеся данные другие формы тренда. Чтобы построить *гиперболический тренд* рассчитаем значения  $1/t$  (табл. 7.5), которые затем будут использоваться в качестве значений объясняющей переменной.

Таблица 7.5

Месяц	$y_t$	$\left(\frac{1}{t}\right)$	$\left(\frac{1}{t}\right)^2$	$y_t \cdot \left(\frac{1}{t}\right)$	$(1/t - \overline{1/t})^2$
1	3,74	1	1	3,74	0,550
2	4,32	0,5	0,25	2,16	0,058
3	4,65	0,333	0,111	1,55	0,006
4	4,79	0,25	0,0625	1,198	0,0001
5	5,01	0,2	0,04	1,002	0,003
6	5,4	0,167	0,028	0,9	0,008
7	5,51	0,143	0,020	0,787	0,013
8	5,88	0,125	0,016	0,735	0,018
9	6,43	0,111	0,012	0,714	0,022
10	6,75	0,1	0,01	0,675	0,025
11	6,99	0,091	0,008	0,635	0,028
12	7,6	0,083	0,007	0,633	0,031
Сумма	67,07	3,103	1,565	14,73	0,762
Среднее	5,589	0,259			

Пользуясь итоговой строкой таблицы 7.5, получаем:

$$\begin{cases} 67,07 = 12b_0 + 3,103b_1 \\ 14,73 = 3,103b_0 + 1,565b_1 \end{cases}$$

$$b_0 = \frac{\Delta_{b_0}}{\Delta} = \frac{59,253}{9,15} = 6,476. \quad b_1 = \frac{\Delta_{b_1}}{\Delta} = \frac{-31,374}{9,15} = -3,429.$$

Тогда уравнение гиперболического тренда запишется следующим образом:

$$y_t = 6,476 - \frac{3,429}{t} + e_t.$$

Выровненные значения  $\hat{y}_t$  по данной модели можно получить, подставляя значения условного времени  $t$  в знаменатель второго слагаемого уравнения тренда либо перемножая значения

1/t на коэффициент регрессии  $b_1$  и прибавляя свободный член  $b_0$ :

$$\hat{y}_1 = 6,476 - \frac{3,429}{1} = 3,047. \quad \hat{y}_2 = 6,476 - \frac{3,429}{2} = 4,761.$$

$$\hat{y}_1 = 6,476 - 3,429 \cdot 1 = 3,047. \quad \hat{y}_2 = 6,476 - 3,429 \cdot 0,5 = 4,761.$$

Таблица 7.6

$t$	$\hat{y}_t$	$(\hat{y}_t - \bar{y})^2$	$(y_t - \hat{y}_t)^2$	$ y_t - \hat{y}_t  / y_t \cdot 100$
1	3,047	6,463	0,480	18,530
2	4,761	0,685	0,195	10,218
3	5,333	0,066	0,466	14,686
4	5,619	0,001	0,687	17,300
5	5,790	0,040	0,609	15,571
6	5,904	0,099	0,254	9,341
7	5,986	0,158	0,227	8,640
8	6,047	0,210	0,028	2,845
9	6,095	0,256	0,112	5,212
10	6,133	0,296	0,381	9,141
11	6,164	0,331	0,682	11,814
12	6,190	0,361	1,988	18,551
<b>Сумма</b>		<b>8,965</b>	<b>6,108</b>	<b>141,848</b>

Пользуясь данными итоговой строки таблицы 7.6, рассчитаем показатели качества гиперболического тренда:

- индекс нелинейной детерминации

$$R^2 = \frac{Q_R}{Q_R + Q_e} = \frac{8,965}{8,965 + 6,108} = 0,595.$$

- средняя ошибка аппроксимации:

$$\bar{A} = \frac{1}{n} \sum \frac{|y_t - \hat{y}_t|}{y_t} \cdot 100\% = \frac{1}{12} \cdot 141,848 = 11,821.$$

Рассчитаем точечный прогноз оборота розничной торговли на три месяца вперед. Для этого подставим в имеющееся уравнение тренда значения  $t_{n+k} = 13, 14, 15$ :

$$\hat{y}_{13} = 6,476 - \frac{3,429}{13} = 6,212.$$

$$\hat{y}_{14} = 6,476 - \frac{3,429}{14} = 6,231.$$

$$\hat{y}_{15} = 6,476 - \frac{3,429}{15} = 6,247.$$

Так как для данного случая  $t_{0,95;10} = 2,23$ , нижние границы доверительных интервалов для соответствующих прогнозных значений будут следующими:

$$\hat{y}_{13} - S_{y_{13}} t_{0,95;10} = 6,212 - 0,782 \cdot \sqrt{1 + \frac{1}{12} + \frac{(1/13 - 0,259)^2}{0,762}} \cdot 2,23 = 4,362.$$

$$\hat{y}_{14} - S_{y_{14}} t_{0,95;10} = 6,231 - 0,782 \cdot \sqrt{1 + \frac{1}{12} + \frac{(1/14 - 0,259)^2}{0,762}} \cdot 2,23 = 4,379.$$

$$\hat{y}_{15} - S_{y_{15}} t_{0,95;10} = 6,247 - 0,782 \cdot \sqrt{1 + \frac{1}{12} + \frac{(1/15 - 0,259)^2}{0,762}} \cdot 2,23 = 4,393,$$

а верхние – следующими:

$$\hat{y}_{13} + S_{y_{13}} t_{0,95;10} = 6,212 + 0,782 \cdot \sqrt{1 + \frac{1}{12} + \frac{(1/13 - 0,259)^2}{0,762}} \cdot 2,23 = 8,062.$$

$$\hat{y}_{14} + S_{y_{14}} t_{0,95;10} = 6,231 + 0,782 \cdot \sqrt{1 + \frac{1}{12} + \frac{(1/14 - 0,259)^2}{0,762}} \cdot 2,23 = 8,083.$$

$$\hat{y}_{15} + S_{y_{15}} t_{0,95;10} = 6,247 + 0,782 \cdot \sqrt{1 + \frac{1}{12} + \frac{(1/15 - 0,259)^2}{0,762}} \cdot 2,23 = 8,101.$$

Представим полученные результаты расчетов на графике (рис. 7.13).

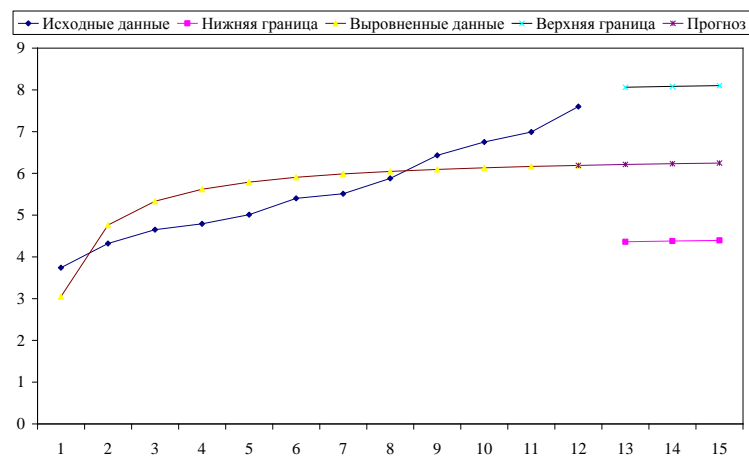


Рисунок 7.13 – Соотношение наблюдаемых  $y_t$ , выровненных  $\hat{y}_t$  и прогнозных  $\hat{y}_{n+k}$  значений для гиперболы

Для построения *полиномиального тренда* в уравнение линейного тренда необходимо добавить переменную  $t^2$  и оценить параметры уравнения множественной регрессии.



Таблица 7.7

Месяц	$y_t$	$t$	$t^2$	$t^3$	$t^4$	$y_t \cdot t$	$y_t \cdot t^2$
1	3,74	1	1	1	1	3,74	3,74
2	4,32	2	4	8	16	8,64	17,28
3	4,65	3	9	27	81	13,95	41,85
4	4,79	4	16	64	256	19,16	76,64
5	5,01	5	25	125	625	25,05	125,25
6	5,4	6	36	216	1296	32,4	194,4
7	5,51	7	49	343	2401	38,57	269,99
8	5,88	8	64	512	4096	47,04	376,32
9	6,43	9	81	729	6561	57,87	520,83
10	6,75	10	100	1000	10000	67,5	675
11	6,99	11	121	1331	14641	76,89	845,79
12	7,6	12	144	1728	20736	91,2	1094,4
<b>Сумма</b>	<b>67,07</b>	<b>78</b>	<b>650</b>	<b>6084</b>	<b>60710</b>	<b>482,01</b>	<b>4241,49</b>

Исходя из данных итоговой строки таблицы 7.7, система уравнений для полинома второй степени примет следующий вид:

$$\begin{cases} 67,07 = 12b_0 + 78b_1 + 650b_2 \\ 482,01 = 78b_0 + 650b_1 + 6084b_2 \\ 4241,49 = 650b_0 + 6084b_1 + 60710b_2 \end{cases},$$

тогда

$$b_0 = \frac{\Delta_{b_0}}{\Delta} = \frac{8517268,8}{2290288} = 3,719. \quad b_1 = \frac{\Delta_{b_1}}{\Delta} = \frac{518626,68}{2290288} = 0,226.$$

$$b_2 = \frac{\Delta_{b_2}}{\Delta} = \frac{16845,4}{2290288} = 0,007.$$

Модель:  $y_t = 3,719 + 0,226x_t + 0,007x_t^2 + e_t.$

Расчет выровненных значений  $\hat{y}_t$  переменной  $Y$  производился путем подстановки  $t$  и  $t^2$  в полученное уравнение тренда:

$$\hat{y}_1 = 3,719 + 0,226 \cdot 1 + 0,007 \cdot 1 = 3,953.$$

$$\hat{y}_2 = 3,719 + 0,226 \cdot 2 + 0,007 \cdot 4 = 4,201.$$

...

$$\hat{y}_{12} = 3,719 + 0,226 \cdot 12 + 0,007 \cdot 144 = 7,495.$$

Таблица 7.8

$t$	$\hat{y}_t$	$(\hat{y}_t - \bar{y})^2$	$(y_t - \hat{y}_t)^2$	$ y_t - \hat{y}_t  / y_t \cdot 100$
1	3,953	2,678	0,045	5,686
2	4,201	1,927	0,014	2,751
3	4,464	1,265	0,034	3,991
4	4,742	0,717	0,002	0,995
5	5,035	0,307	0,001	0,498
6	5,342	0,061	0,003	1,068
7	5,664	0,006	0,024	2,802
8	6,001	0,170	0,015	2,061
9	6,353	0,583	0,006	1,203
10	6,719	1,276	0,001	0,462

11	7,100	2,282	0,012	1,570
12	7,495	3,634	0,011	1,377
<b>Сумма</b>		<b>14,905</b>	<b>0,168</b>	<b>24,464</b>

Пользуясь последней строкой таблицы 7.8, рассчитаем показатели качества полученной модели:

- индекс нелинейной детерминации

$$R^2 = \frac{Q_R}{Q_R + Q_e} = \frac{14,905}{14,905 + 0,168} = 0,989.$$

- средняя ошибка аппроксимации

$$\bar{A} = \frac{1}{n} \sum \frac{|y_t - \hat{y}_t|}{y_t} \cdot 100\% = \frac{1}{12} \cdot 24,464 = 2,039.$$

Рассчитаем точечный прогноз оборота розничной торговли на три месяца вперед. Для этого подставим в имеющееся уравнение тренда значения  $t_{n+k} = 13, 14, 15$  и их квадраты  $t_{n+k}^2 = 169, 196, 225$ :

$$\hat{y}_{13} = 3,719 + 0,226 \cdot 13 + 0,007 \cdot 169 = 7,906.$$

$$\hat{y}_{14} = 3,719 + 0,226 \cdot 14 + 0,007 \cdot 196 = 8,331.$$

$$\hat{y}_{15} = 3,719 + 0,226 \cdot 15 + 0,007 \cdot 225 = 8,77.$$

Для получения интервального прогноза рассчитаем стандартные ошибки имеющегося точечного прогноза. Их расчет будем производить по следующей формуле:

$$S_{y_{n+k}} = S_e \sqrt{1 + T_p^T (T^T T)^{-1} T_p},$$

где

$$T_p^T = (1 \quad t \quad t^2), \quad T_p = \begin{pmatrix} 1 \\ t \\ t^2 \end{pmatrix}.$$

$(T^T T)^{-1}$  – матрица, обратная основной матрице системы линейных уравнений:

$$(T^T T) = \begin{pmatrix} 12 & 78 & 650 \\ 78 & 650 & 6084 \\ 650 & 6084 & 60710 \end{pmatrix}.$$

Для данного примера обратная матрица  $(T^T T)^{-1}$  примет следующий вид:

$$(T^T T)^{-1} = \begin{pmatrix} 1,06818 & -0,34091 & 0,02273 \\ -0,34091 & 0,13362 & -0,00974 \\ 0,02273 & -0,00974 & 0,00075 \end{pmatrix}.$$

Тогда стандартные ошибки будут следующими:

$$S_{y_{13}} = 0,137 \sqrt{1 + (1 \ 13 \ 169) \begin{pmatrix} 1,06818 & -0,34091 & 0,02273 \\ -0,34091 & 0,13362 & -0,00974 \\ 0,02273 & -0,00974 & 0,00075 \end{pmatrix} \begin{pmatrix} 1 \\ 13 \\ 169 \end{pmatrix}} = 0,197.$$

$$S_{y_{14}} = 0,137 \sqrt{1 + (1 \ 14 \ 196) \begin{pmatrix} 1,06818 & -0,34091 & 0,02273 \\ -0,34091 & 0,13362 & -0,00974 \\ 0,02273 & -0,00974 & 0,00075 \end{pmatrix} \begin{pmatrix} 1 \\ 14 \\ 196 \end{pmatrix}} = 0,235.$$

$$S_{y_{15}} = 0,137 \sqrt{1 + (1 \ 15 \ 225) \begin{pmatrix} 1,06818 & -0,34091 & 0,02273 \\ -0,34091 & 0,13362 & -0,00974 \\ 0,02273 & -0,00974 & 0,00075 \end{pmatrix} \begin{pmatrix} 1 \\ 15 \\ 225 \end{pmatrix}} = 0,284.$$

Так как для данного случая  $t_{0,95;9} = 2,26$ , нижние границы доверительных интервалов для соответствующих прогнозных значений будут следующими:

$$\hat{y}_{13} - S_{y_{13}} t_{0,95;9} = 7,906 - 0,197 \cdot 2,26 = 7,461.$$

$$\hat{y}_{14} - S_{y_{14}} t_{0,95;9} = 8,331 - 0,235 \cdot 2,26 = 7,8.$$

$$\hat{y}_{15} - S_{y_{15}} t_{0,95;9} = 8,77 - 0,284 \cdot 2,26 = 8,128,$$

а верхние – следующими:

$$\hat{y}_{13} + S_{y_{13}} t_{0,95;9} = 7,906 + 0,197 \cdot 2,26 = 8,35.$$

$$\hat{y}_{14} + S_{y_{14}} t_{0,95;9} = 8,331 + 0,235 \cdot 2,26 = 8,862.$$

$$\hat{y}_{15} + S_{y_{15}} t_{0,95;9} = 8,77 + 0,284 \cdot 2,26 = 9,413.$$

Полученные результаты расчетов представим на графике (рис. 7.14).

Для построения *степенного тренда* вместо исходных данных  $y_t$  и  $t$  будем использовать их логарифмы  $\ln y_t$  и  $\ln t$ . Пользуясь данными таблицы 7.9, получаем:

$$\begin{cases} 20,405 = 12 \ln b_0 + b_1 19,987 \\ 35,682 = 19,987 \ln b_0 + 39,575 b_1 \end{cases}$$

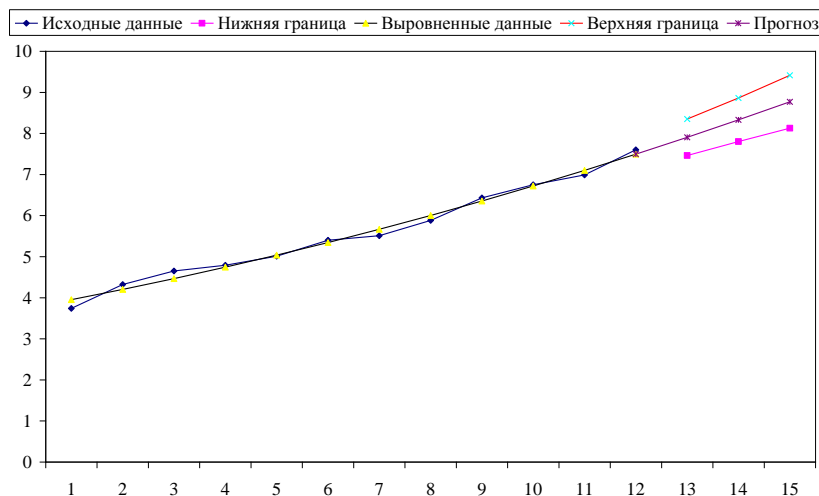


Рисунок 7.14 – Соотношение наблюдаемых  $y_t$ , выровненных  $\hat{y}_t$

и прогнозных  $\hat{y}_{n+k}$  значений для полинома 2-й степени

Таблица 7.9

$y_t$	$t$	$\ln y_t$	$\ln t$	$(\ln t)^2$	$\ln y_t \ln t$	$(t - \bar{t})^2$
3,74	1	1,319	0	0	0	30,25
4,32	2	1,463	0,693	0,480	1,014	20,25
4,65	3	1,537	1,099	1,207	1,688	12,25
4,79	4	1,567	1,386	1,922	2,172	6,25
5,01	5	1,611	1,609	2,590	2,594	2,25
5,40	6	1,686	1,792	3,210	3,022	0,25
5,51	7	1,707	1,946	3,787	3,321	0,25
5,88	8	1,772	2,079	4,324	3,684	2,25
6,43	9	1,861	2,197	4,828	4,089	6,25
6,75	10	1,910	2,303	5,302	4,397	12,25
6,99	11	1,944	2,398	5,750	4,663	20,25
7,60	12	2,028	2,485	6,175	5,040	30,25
<b>Сумма</b>		<b>20,405</b>	<b>19,987</b>	<b>39,575</b>	<b>35,682</b>	<b>143</b>
<b>Среднее</b>		<b>1,700</b>				

Решая систему по правилу Крамера, получаем:

$$\ln b_0 = \frac{\Delta_{\ln b_0}}{\Delta} = \frac{94,327}{75,41} = 1,251. \quad b_1 = \frac{\Delta_{b_1}}{\Delta} = \frac{20,353}{75,41} = 0,27.$$

Тогда степенной тренд в линеаризованной форме примет следующий вид:

$$\ln y_t = 1,251 + 0,271 \cdot \ln t + \ln e_t.$$

Преобразуем полученную модель к нелинейному виду. Для этого воспользуемся числом  $e = 2,71828$ :

$$y_t = 2,71828^{1,251} \cdot t^{0,271} \cdot e_t \Rightarrow \\ \Rightarrow y_t = 3,493 \cdot t^{0,271} \cdot e_t.$$

Выровненные значения  $\hat{y}_t$  по данной модели можно получить двумя способами. Первый заключается в том, что в линеаризованное уравнение тренда подставляются значения логарифмов условного времени  $\ln t$ , то есть:

$$\ln \hat{y}_t = 1,251 + 0,271 \cdot \ln t$$

$$\ln \hat{y}_1 = 1,251 + 0,271 \cdot 0 = 1,251.$$

$$\ln \hat{y}_2 = 1,251 + 0,271 \cdot 0,693 = 1,438.$$

...

$$\ln \hat{y}_{12} = 1,251 + 0,271 \cdot 2,485 = 1,922.$$

После чего полученные логарифмы выровненных значений  $\ln \hat{y}_t$  приводят к исходным единицам измерения:

$$\hat{y}_1 = e^{\ln \hat{y}_1} = 2,71828^{1,251} = 3,493.$$

$$\hat{y}_2 = e^{\ln \hat{y}_2} = 2,71828^{1,438} = 4,212.$$

...

$$\hat{y}_{12} = e^{\ln \hat{y}_{12}} = 2,71828^{1,922} = 6,831$$

Второй способ заключается в том, что в нелинейный тренд подставляются исходные значения условного времени  $t$ , то есть

$$\hat{y}_t = 3,493 \cdot t^{0,271}$$

$$\hat{y}_1 = 3,493 \cdot 1^{0,271} = 3,493.$$

$$\hat{y}_2 = 3,493 \cdot 2^{0,271} = 4,212.$$

...

$$\hat{y}_{12} = 3,493 \cdot 12^{0,271} = 6,831.$$

В нашем примере воспользуемся первым способом (табл. 7.10)

Пользуясь данными итоговой строки таблицы 7.10, рассчитаем показатели качества степенного тренда:

- индекс нелинейной детерминации

$$R^2 = \frac{Q_R}{Q_R + Q_e} = \frac{15,073}{15,073 + 1,356} = 0,917.$$

- средняя ошибка аппроксимации

$$\bar{A} = \frac{1}{n} \sum \frac{|y_t - \hat{y}_t|}{y_t} \cdot 100\% = \frac{1}{12} \cdot 60,089 = 5,007.$$

Таблица 7.10

$t$	$\ln \hat{y}_t$	$\hat{y}_t$	$(\hat{y}_t - \bar{y})^2$	$(y_t - \hat{y}_t)^2$	$ y_t - \hat{y}_t  / y_t \cdot 100$
1	1,251	3,493	3,419	0,061	6,595
2	1,438	4,212	1,611	0,012	2,500
3	1,547	4,699	0,882	0,002	1,056
4	1,625	5,079	0,639	0,083	6,023
5	1,685	5,394	0,335	0,147	7,660
6	1,734	5,666	0,036	0,071	4,923
7	1,776	5,907	0,006	0,157	7,197
8	1,812	6,123	0,085	0,059	4,138
9	1,844	6,321	0,707	0,012	1,694
10	1,872	6,503	1,348	0,061	3,653
11	1,898	6,673	1,962	0,101	4,537
12	1,922	6,831	4,043	0,591	10,113
<b>Сумма</b>			<b>15,073</b>	<b>1,356</b>	<b>60,089</b>

Рассчитаем прогноз оборота розничной торговли на три месяца вперед. Для этого в линеаризованное уравнение

$$\ln \hat{y}_t = 1,251 + 0,271 \cdot \ln t$$

подставим логарифмированные значения условного времени  $\ln t = 2,565; 2,639; 2,708$ , соответствующие исходным значениям  $t = 13; 14; 15$ :

$$\ln \hat{y}_{13} = 1,251 + 0,271 \cdot 2,565 = 1,94.$$

$$\ln \hat{y}_{14} = 1,251 + 0,271 \cdot 2,639 = 1,96.$$

$$\ln \hat{y}_{15} = 1,251 + 0,271 \cdot 2,708 = 1,98.$$

Приведем полученные значения  $\ln \hat{y}_t$  к исходным единицам измерения:

$$\hat{y}_{13} = e^{\ln \hat{y}_{13}} = 2,71828^{1,94} = 6,96.$$

$$\hat{y}_{14} = e^{\ln \hat{y}_{14}} = 2,71828^{1,96} = 7,1$$

$$\hat{y}_{15} = e^{\ln \hat{y}_{15}} = 2,71828^{1,98} = 7,24$$

Так как для данного случая  $t_{0,95;10} = 2,23$ , нижние границы доверительных интервалов для соответствующих прогнозных значений будут следующими:

$$\hat{y}_{13} - S_{y_{13}} \cdot t_{0,95;10} = 6,96 - 0,368 \cdot \sqrt{1 + \frac{1}{12} + \frac{(13-6,5)^2}{143}} \cdot 2,23 = 6,02.$$

$$\hat{y}_{14} - S_{y_{14}} \cdot t_{0,95;10} = 7,1 - 0,368 \cdot \sqrt{1 + \frac{1}{12} + \frac{(14-6,5)^2}{143}} \cdot 2,23 = 6,124.$$

$$\hat{y}_{15} - S_{y_{15}} \cdot t_{0,95;10} = 7,24 - 0,368 \cdot \sqrt{1 + \frac{1}{12} + \frac{(15-6,5)^2}{143}} \cdot 2,23 = 6,22,$$

а верхние – следующими:

$$\hat{y}_{13} + S_{y_{13}} \cdot t_{0,95;10} = 6,96 + 0,368 \cdot \sqrt{1 + \frac{1}{12} + \frac{(13-6,5)^2}{143}} \cdot 2,23 = 7,945.$$

$$\hat{y}_{14} + S_{y_{14}} \cdot t_{0,95;10} = 7,1 + 0,368 \cdot \sqrt{1 + \frac{1}{12} + \frac{(14 - 6,5)^2}{143}} \cdot 2,23 = 8,12.$$

$$\hat{y}_{15} + S_{y_{15}} \cdot t_{0,95;10} = 7,24 + 0,368 \cdot \sqrt{1 + \frac{1}{12} + \frac{(15 - 6,5)^2}{143}} \cdot 2,23 = 8,29.$$

Полученные результаты расчетов представим на графике (рис. 7.15).

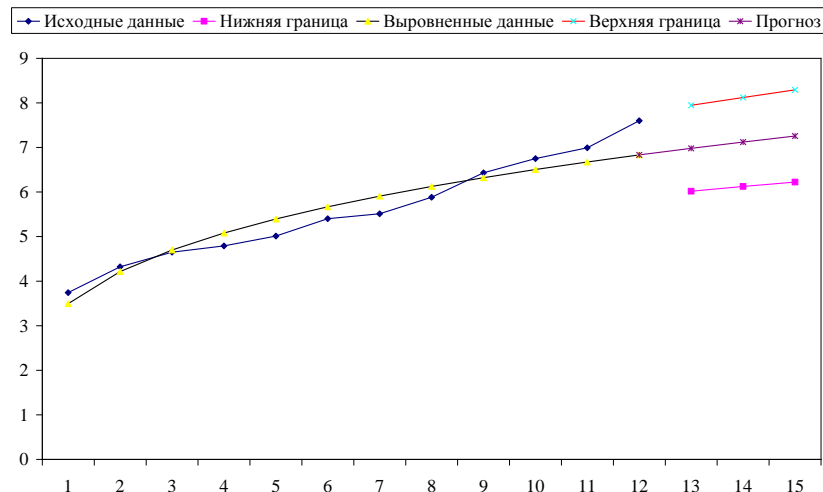


Рисунок 7.15 – Соотношение наблюдаемых  $y_t$ , выровненных  $\hat{y}_t$  и прогнозных  $\hat{y}_{n+k}$  значений для степенной функции

Далее построим уравнение тренда с использованием *показательной функции*. Для оценки его параметров будем использовать логарифмы наблюдаемых значений временного ряда  $\ln y_t$  и исходные значения условного времени  $t$ .

Пользуясь данными таблицы 7.11, получаем следующую систему уравнений:

$$\begin{cases} 20,405 = 12 \ln b_0 + 78 \ln b_1 \\ 140,988 = 78 \ln b_0 + 650 \ln b_1 \end{cases}$$

$$\ln b_0 = \frac{\Delta_{\ln b_0}}{\Delta} = \frac{2266,117}{1716} = 1,321. \quad \ln b_1 = \frac{\Delta_{\ln b_1}}{\Delta} = \frac{100,273}{1716} = 0,058.$$

Степенной тренд в линеаризованной форме примет следующий вид:

$$\ln y_t = 1,321 + 0,058 \cdot t + \ln e_t.$$

Преобразуем полученную модель к нелинейному виду. Для этого воспользуемся числом  $e = 2,71828$ :

$$\begin{aligned} y_t &= e^{1,321} \cdot (e^{0,058})^t \cdot e_t \Rightarrow \\ \Rightarrow y_t &= 2,71828^{1,321} \cdot (2,71828^{0,058})^t \cdot e_t \Rightarrow \\ \Rightarrow y_t &= 3,746 \cdot 1,06^t \cdot e_t. \end{aligned}$$

Таблица 7.11

Месяц	$y_t$	$\ln y_t$	$t$	$t^2$	$t \cdot \ln y_t$	$(t - \bar{t})^2$
1	3,74	1,319	1	1	1,319	30,25
2	4,32	1,463	2	4	2,927	20,25
3	4,65	1,537	3	9	4,611	12,25
4	4,79	1,567	4	16	6,266	6,25

Месяц	$y_t$	$\ln y_t$	$t$	$t^2$	$t \cdot \ln y_t$	$(t - \bar{t})^2$
5	5,01	1,611	5	25	8,057	2,25
6	5,4	1,686	6	36	10,118	0,25
7	5,51	1,707	7	49	11,946	0,25
8	5,88	1,772	8	64	14,172	2,25
9	6,43	1,861	9	81	16,749	6,25
10	6,75	1,910	10	100	19,095	12,25
11	6,99	1,944	11	121	21,389	20,25
12	7,6	2,028	12	144	24,338	30,25
<b>Сумма</b>	<b>20,405</b>	<b>78</b>	<b>650</b>	<b>140,988</b>	<b>143</b>	
<b>Среднее</b>	<b>1,700</b>	<b>6,5</b>				

Расчет логарифмов выровненных значений  $\ln \hat{y}_t$  моделируемого временного ряда будем производить путем подстановки наблюдаемых значений условного времени  $t$  в линейаризованное уравнение тренда (табл. 7.12).

$$\ln \hat{y}_t = 1,321 + 0,058 \cdot t.$$

$$\ln \hat{y}_1 = 1,321 + 0,058 \cdot 1 = 1,379.$$

$$\ln \hat{y}_2 = 1,321 + 0,058 \cdot 2 = 1,437.$$

...

$$\ln \hat{y}_{12} = 1,321 + 0,058 \cdot 12 = 2,02$$

Таблица 7.12

$t$	$\ln \hat{y}_t$	$\hat{y}_t$	$(\hat{y}_t - \bar{y})^2$	$(y_t - \hat{y}_t)^2$	$ y_t - \hat{y}_t  / y_t \cdot 100$
1	1,379	3,971	2,618	0,0534	6,176
2	1,437	4,210	1,902	0,0121	2,548
3	1,496	4,463	1,268	0,0349	4,016
4	1,554	4,732	0,735	0,0034	1,214
5	1,613	5,017	0,328	0,0000	0,132
6	1,671	5,318	0,073	0,0066	1,510
7	1,730	5,639	0,002	0,0165	2,332
8	1,788	5,978	0,151	0,0096	1,663
9	1,846	6,338	0,560	0,0086	1,438
10	1,905	6,719	1,276	0,0010	0,461
11	1,963	7,123	2,353	0,0177	1,906
12	2,022	7,552	3,852	0,0023	0,634
<b>Сумма</b>			<b>15,120</b>	<b>0,166</b>	<b>24,029</b>

Приведем логарифмы выровненных значений  $\ln \hat{y}_t$  к исходным единицам измерения  $\hat{y}_t$ :

$$\hat{y}_1 = e^{\ln \hat{y}_1} = 2,71828^{1,379} = 3,97.$$

$$\hat{y}_2 = e^{\ln \hat{y}_2} = 2,71828^{1,437} = 4,21.$$

...

$$\hat{y}_{12} = e^{\ln \hat{y}_{12}} = 2,71828^{2,02} = 7,55.$$

Пользуясь данными итоговой строки таблицы 7.12, рассчитаем показатели качества показательного тренда:

- индекс нелинейной детерминации

$$R^2 = \frac{Q_R}{Q_R + Q_e} = \frac{15,12}{15,12 + 0,166} = 0,989.$$

- средняя ошибка аппроксимации



$$\bar{A} = \frac{1}{n} \sum \frac{|y_t - \hat{y}_t|}{y_t} \cdot 100\% = \frac{1}{12} \cdot 24,029 = 2,002.$$

Рассчитаем прогноз оборота розничной торговли на три месяца вперед. Для этого в линеаризованное уравнение

$$\ln \hat{y}_t = 1,321 + 0,058 \cdot t$$

подставим значения условного времени  $t = 13; 14; 15$ :

$$\ln \hat{y}_{13} = 1,321 + 0,058 \cdot 13 = 2,08.$$

$$\ln \hat{y}_{14} = 1,321 + 0,058 \cdot 14 = 2,13.$$

$$\ln \hat{y}_{15} = 1,321 + 0,058 \cdot 15 = 2,19.$$

Приведем полученные прогнозные значения  $\ln \hat{y}_t$  к исходным единицам измерения:

$$\hat{y}_{13} = e^{\ln \hat{y}_{13}} = 2,71828^{2,08} = 8,004.$$

$$\hat{y}_{14} = e^{\ln \hat{y}_{14}} = 2,71828^{2,13} = 8,42.$$

$$\hat{y}_{15} = e^{\ln \hat{y}_{15}} = 2,71828^{2,19} = 8,94.$$

Так как для данного случая  $t_{0,95;10} = 2,23$ , нижние границы доверительных интервалов для соответствующих прогнозных значений будут следующими:

$$\hat{y}_{13} - S_{y_{13}} \cdot t_{0,95;10} = 8,004 - 0,129 \cdot \sqrt{1 + \frac{1}{12} + \frac{(13-6,5)^2}{143}} \cdot 2,23 = 7,67.$$

$$\hat{y}_{14} - S_{y_{14}} \cdot t_{0,95;10} = 8,42 - 0,129 \cdot \sqrt{1 + \frac{1}{12} + \frac{(14-6,5)^2}{143}} \cdot 2,23 = 8,1.$$

$$\hat{y}_{15} - S_{y_{15}} \cdot t_{0,95;10} = 8,94 - 0,129 \cdot \sqrt{1 + \frac{1}{12} + \frac{(15-6,5)^2}{143}} \cdot 2,23 = 8,6,$$

а верхние – следующими:

$$\hat{y}_{13} + S_{y_{13}} \cdot t_{0,95;10} = 8,004 + 0,129 \cdot \sqrt{1 + \frac{1}{12} + \frac{(13-6,5)^2}{143}} \cdot 2,23 = 8,34.$$

$$\hat{y}_{14} + S_{y_{14}} \cdot t_{0,95;10} = 8,42 + 0,129 \cdot \sqrt{1 + \frac{1}{12} + \frac{(14-6,5)^2}{143}} \cdot 2,23 = 8,8.$$

$$\hat{y}_{15} + S_{y_{15}} \cdot t_{0,95;10} = 8,94 + 0,129 \cdot \sqrt{1 + \frac{1}{12} + \frac{(15-6,5)^2}{143}} \cdot 2,23 = 9,3.$$

Полученные результаты расчетов представим на графике (рис. 7.16).

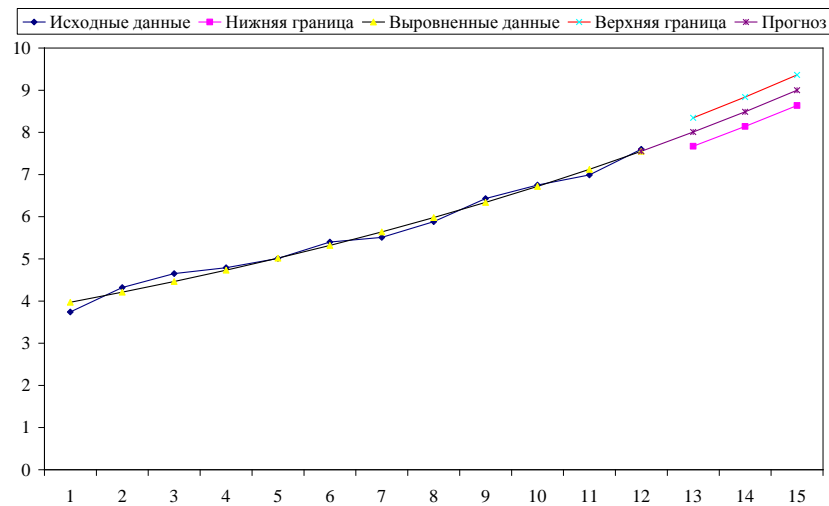


Рисунок 7.16 – Соотношение наблюдаемых  $y_t$ , выровненных  $\hat{y}_t$  и прогнозных  $\hat{y}_{n+k}$  значений для показательной функции

3. Как уже отмечалось, выбор формы тренда, наилучшим образом описывающего имеющийся временной ряд, производится на основе показателей качества. От качества построенного тренда зависит ширина доверительного интервала выполненного прогноза. Чем выше качество, тем уже интервал. От ширины интервала зависит вероятность реализации в будущем выполненного точечного прогноза. Чем уже интервал, тем выше эта вероятность, тем меньшему риску мы подвергаемся, руководствуясь значениями выполненного прогноза.

Представим в таблице основные результаты выполненных расчетов (табл. 7.13).

Таблица 7.13

Функция	Линейная	Гипербола	Парабола	Степенная	Показательная
$t_{n+k}$	$R_{y_{t+\tau}}$	$R_{y_{t+\tau}}$	$R_{y_{t+\tau}}$	$R_{y_{t+\tau}}$	$R_{y_{t+\tau}}$
13	0,812	3,700	0,889	1,929	0,675
14	0,841	3,704	1,062	1,996	0,698
15	0,872	3,708	1,285	2,070	0,724
Сумма	2,525	11,112	3,236	5,995	2,098
Среднее	0,842	3,704	1,079	1,998	0,699
$R^2$	0,984	0,595	0,9888	0,917	0,9891
$\bar{A}$	2,115	11,821	2,039	5,007	2,002

где  $R_{y_{t+\tau}}$  – размах доверительного интервала для точечного прогноза  $\hat{y}_{n+k}$  (разница между верхней и нижней границами)

Из таблицы видно, что минимальный средний размах доверительного интервала  $\bar{R}_{y_{t+\tau}} = 0,699$ , минимальная средняя ошибка аппроксимации  $\bar{A} = 2,002$  и максимальный индекс детерминации  $R^2 = 0,9891$  соответствуют показательной функции. Можно сделать вывод о том, что показательный тренд наилучшим образом описывает имеющийся временной ряд. Следовательно, для прогнозирования оборота розничной торговли целесообразнее использовать именно показательный тренд.

При прогнозировании любого временного ряда всегда делается важное допущение о том, что характер динамики изучаемого процесса  $Y$ , наблюдаемый на исследуемом периоде, сохранится и не будет меняться в будущем. Однако иногда приходится сталкиваться со случаями, когда известно, что в будущем динамика моделируемого показателя замедлится либо ускорится, не смотря на иной ее характер в исследуемом периоде. Зная такую информацию, исследователь, невзирая на качество тренда, может подобрать такую кривую, которая будет наилучшим образом соответствовать его субъективным предположениям о будущей динамике процесса.

Представим на графике наряду с имеющимся временным рядом все имеющиеся варианты прогнозов (рис. 7.17).

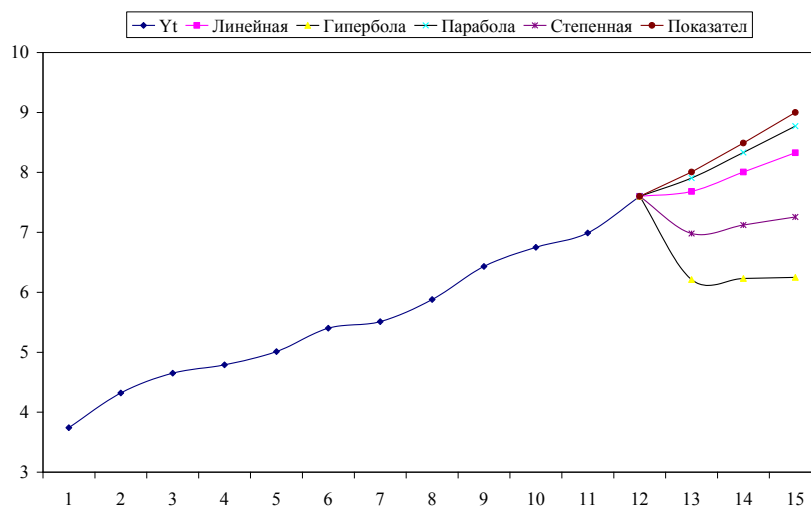


Рисунок 7.17 – Прогнозы по пяти построенным трендам

Если учесть то, что конец 2008 г. соответствует началу мирового финансово-экономического кризиса, закономерно предположить, что рост оборота розничной торговли будет уже не столь интенсивен, и для его прогнозирования в данной ситуации предпочтительней воспользоваться степенным трендом.

### 7.7. Моделирование сезонных колебаний

Как уже было отмечено, наряду с моделированием тенденции одной из важнейших задач анализа экономических временных рядов также является *моделирование сезонных колебаний*. Смоделировать только тенденцию не всегда достаточно, так как трендовая модель, ввиду высокой колеблемости моделируемого показателя, может иметь низкое качество, а полученные прогнозные значения будут далеки от действительности. Для повышения качества модели, а это является залогом более точного прогноза, прибегают к моделированию сезонных колебаний.

Необходимо так же понимать, что наличие высокой колеблемости во временном ряду не всегда свидетельствует о том, что моделирование сезонных колебаний улучшит модель. Данную методику целесообразно применять лишь тогда, когда *доказано*, что одинаковые по характеру и направленности колебания показателя повторяются с равной периодичностью (например, в течение каждого года).

Подобные колебания можно идентифицировать визуально, то есть на графике исходных данных, или количественно – на основе рассчитанных значений *коэффициентов автокорреляции* порядка  $\tau$ . Если наиболее высоким оказался коэффициент автокорреляции порядка  $\tau$  (отличный от единицы), то временной ряд содержит сезонные колебания с периодичностью в  $\tau$  моментов времени. Следовательно, в данном случае допустимо применение методики моделирования сезонных колебаний.

Существуют две модели сезонных колебаний – *аддитивная* и *мультипликативная*. Выбор одной из них зависит от характера амплитуды сезонных колебаний.

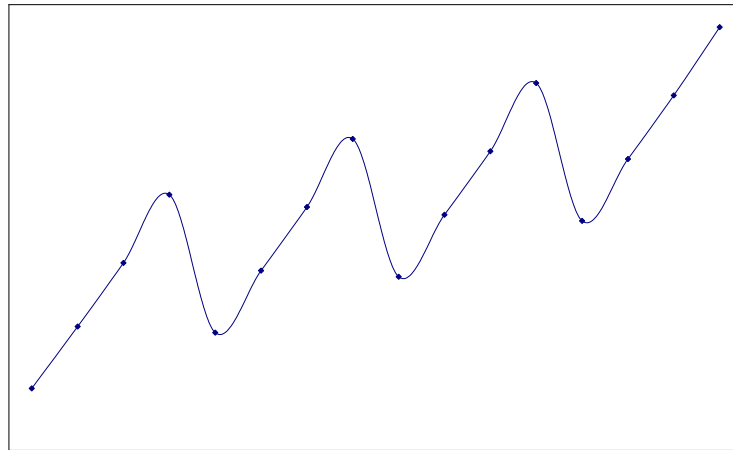


Рисунок 7.18 – Постоянная амплитуда сезонных колебаний

Если амплитуда сезонных колебаний примерно постоянна и с наступлением каждого следующего периода (НЕ МОМЕНТА ВРЕМЕНИ) не демонстрирует тенденций к увеличению или уменьшению (рис. 7.18), то целесообразно использовать аддитивную модель временного ряда. Общий вид аддитивной модели следующий:

$$Y = T + S + E$$

то есть каждый уровень временного ряда  $Y$  здесь представляется в виде суммы трендовой  $T$ , сезонной  $S$  и случайной  $E$  компонент.

Если амплитуда колебаний с наступлением каждого следующего периода уменьшается или возрастает (рис. 7.19), то используется мультипликативная модель. Общий вид мультипликативной модели следующий:

$$Y = T \cdot S \cdot E$$

то есть каждый уровень временного ряда  $Y$  представляется как произведение трендовой  $T$ , сезонной  $S$  и случайной  $E$  компонент.

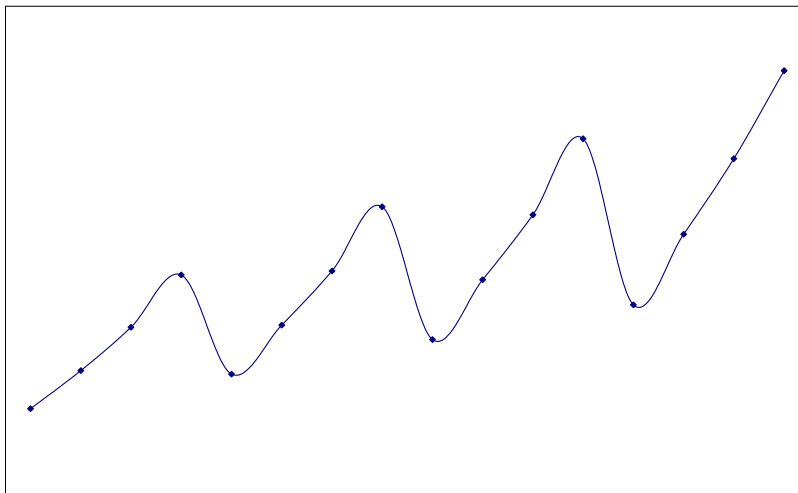


Рисунок 7.19 – Растущая амплитуда сезонных колебаний

Построение аддитивной или мультипликативной моделей сводится к *расчету значений  $T$*  (тенденции),  *$S$*  (сезонных колебаний) и  *$E$*  (случайных колебаний) для каждого уровня ряда. Процесс построения модели можно разбить на шесть шагов.

**1 шаг.** Выравнивание исходного временного ряда  $y_t$  с помощью скользящей средней. Важным моментом здесь является выбор величины *окна усреднения*, зависящей от периодичности моделируемого показателя. Если сезонные колебания повторяются с периодичностью в 4 момента

времени (поквартальные наблюдения), то, соответственно, окно усреднения будет равно четырем, а скользящая средняя будет следующей:

$$\bar{y}_t^c = \frac{y_{t-2} + y_{t-1} + y_t + y_{t+1}}{4}.$$

Если сезонные колебания повторяются с периодичностью в 12 моментов времени (помесячные наблюдения), то окно усреднения будет равно 12, а скользящая средняя будет следующей:

$$\bar{y}_t^c = \frac{y_{t-5} + y_{t-4} + y_{t-3} + y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2} + y_{t+3} + y_{t+4} + y_{t+5} + y_{t+6}}{12}.$$

**2 шаг.** Расчет значений сезонной компоненты  $S$ . Процесс расчетов на данном и последующих этапах будет зависеть от того, какая применяется модель (аддитивная или мультипликативная). Если строится *аддитивная* модель, то на данном шаге выполняется *вычитание* значений скользящей средней  $\bar{y}_t^c$  из соответствующих наблюдаемых значений  $y_t$  моделируемого показателя  $Y$ . Затем получают значения сезонной компоненты  $\bar{S}_i$ , путем усреднения разностей  $(y_t - \bar{y}_t^c)$  поквартально или ежемесячно в зависимости от характера исходных данных  $y_t$ .

Далее полученные коэффициенты сезонности  $\bar{S}_i$  проверяют на взаимопогашаемость. *Свойство взаимопогашаемости сезонных колебаний для аддитивной модели заключается в том, что их сумма должна быть равна нулю.*

Если  $\sum \bar{S}_i \neq 0$ , полученные значения сезонной компоненты корректируются на величину:

$$k = \frac{\sum \bar{S}_i}{\tau},$$

в результате чего получают скорректированные значения коэффициентов сезонности:

$$\hat{S}_i = \bar{S}_i - k.$$

Если строится мультипликативная модель, то на данном шаге выполняется деление наблюдаемых значений  $y_t$  на соответствующие значения скользящей средней  $\bar{y}_t^c$ . Затем так же получают оценки сезонной компоненты  $\bar{S}_i$  усредняя значения  $\frac{y_t}{\bar{y}_t^c}$  поквартально или ежемесячно.

*Свойство взаимопогашаемости сезонных колебаний для мультипликативной модели заключается в том, что их сумма должна быть равна периодичности  $\tau$ .*

Если  $\sum \bar{S}_i \neq \tau$ , полученные значения сезонной компоненты корректируются на величину:

$$k = \frac{\tau}{\sum \bar{S}_i},$$

в результате чего получают скорректированные значения сезонной компоненты:

$$\hat{S}_i = \bar{S}_i \cdot k.$$

**3 шаг.** Устранение сезонной компоненты из исходных уровней ряда. Здесь в случае аддитивной модели из исходных уровней ряда  $y_t$  отнимают соответствующие им значения коэффициентов сезонности  $\hat{S}_i$ . В результате получают выровненные значения  $(T+E)$ , включающие в себя только тренд и случайные колебания. В случае мультипликативной модели – исходные уровни ряда  $y_t$  делят на соответствующие им коэффициенты сезонности  $\hat{S}_i$ . В результате получают выровненные значения  $(T \cdot E)$ , также содержащие тренд и случайные колебания.

**4 шаг.** Выделение чистой тенденции  $T$ . Выделение тенденции производится с помощью аналитического выравнивания уровней  $(T+E)$  для аддитивной модели и уровней  $(T \cdot E)$  для мультипликативной модели. В результате полученные с помощью уравнения тренда выровненные значения служат оценкой чистой тенденции  $T$ .

**5 шаг.** Расчет теоретических значений моделируемого показателя. В случае аддитивной модели они будут рассчитываться как  $\hat{y}_t = T + S$ , а в случае мультипликативной – как  $\hat{y}_t = T \cdot S$  (где  $S = \hat{S}_i$ ).

**6 шаг.** Расчет абсолютных или относительных значений остатков  $E$ . Абсолютные значения  $e_t = y_t - \hat{y}_t$  – для аддитивной модели, относительные  $e_t = y_t / \hat{y}_t$  – для мультипликативной.

Таким образом, в общем виде модели сезонных колебаний можно представить следующим образом:

- аддитивная:  $y_t = T + S + E = (b_0 + b_1 t) + \hat{S}_i + e_t$ .

- мультипликативная:  $y_t = T \cdot S \cdot E = (b_0 + b_1 t) \cdot \hat{S}_i \cdot e_t$ .

## 7.8. Проверка качества модели

Насколько точно построенные модели описывают имеющийся временной ряд, можно проверить с помощью стандартных показателей качества:

коэффициента детерминации

$$R^2 = \frac{\sum (\hat{y}_t - \bar{y}_t)^2}{\sum (\hat{y}_t - \bar{y}_t)^2 + \sum (y_t - \hat{y}_t)^2}$$

и средней ошибки аппроксимации:

$$\bar{A} = \frac{1}{n} \sum \frac{|y_t - \hat{y}_t|}{y_t} \cdot 100\%$$

Однако следует отметить, что вопреки представленной формулировке мультипликативной модели  $y_t = T \cdot S \cdot E$  для расчета показателей качества (коэффициента детерминации и средней ошибки аппроксимации) используются не относительные остатки  $e_t = y_t / \hat{y}_t$ , а абсолютные  $e_t = y_t - T \cdot S$ .

## 7.9. Прогнозирование с помощью моделей сезонных колебаний

Построенная модель сезонных колебаний позволяет выполнить точечный и интервальный прогноз временного ряда. Для этого необходимо продлить переменную  $t$ , используемую при аналитическом выравнивании временных рядов  $(T+E)$  и  $(T \cdot E)$ , на  $k$  моментов времени вперед и подставить полученные значения в трендовую часть аддитивной или мультипликативной модели  $(b_0 + b_1 t)$ . Данная операция позволяет получить прогноз тенденции  $\hat{T}_{n+k}$ . Для отражения сезонных колебаний его необходимо соответствующим образом скорректировать (прибавить  $\hat{S}_i$  или умножить на  $\hat{S}_i$ ) на величину коэффициентов сезонности. Точечный прогноз можно получить следующим образом:

- аддитивная модель  $\hat{y}_{n+k} = \hat{T}_{n+k} + \hat{S}_i = (b_0 + b_1 t_{n+k}) + \hat{S}_i$ .

- мультипликативная модель  $\hat{y}_{n+k} = \hat{T}_{n+k} \cdot \hat{S}_i = (b_0 + b_1 t_{n+k}) \cdot \hat{S}_i$ .

Для получения интервального прогноза необходимо найти верхние и нижние границы доверительных интервалов прогноза:

$$\hat{y}_{n+k} - S_{\hat{y}_{n+k}} \cdot t_{\alpha; df} \leq \hat{y}_{n+k} \leq \hat{y}_{n+k} + S_{\hat{y}_{n+k}} \cdot t_{\alpha; df},$$

где

$$S_{\hat{y}_{n+k}} = S_e \sqrt{1 + \frac{1}{n} + \frac{(t_{n+k} - \bar{t})^2}{\sum (t - \bar{t})^2}} - \text{стандартная ошибка прогноза.}$$

$$S_e = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{n - p - 1}} - \text{стандартная ошибка регрессии.}$$