

ТЕМА: ПОСТРОЕНИЕ АДДИТИВНОЙ МОДЕЛИ ВРЕМЕННОГО РЯДА

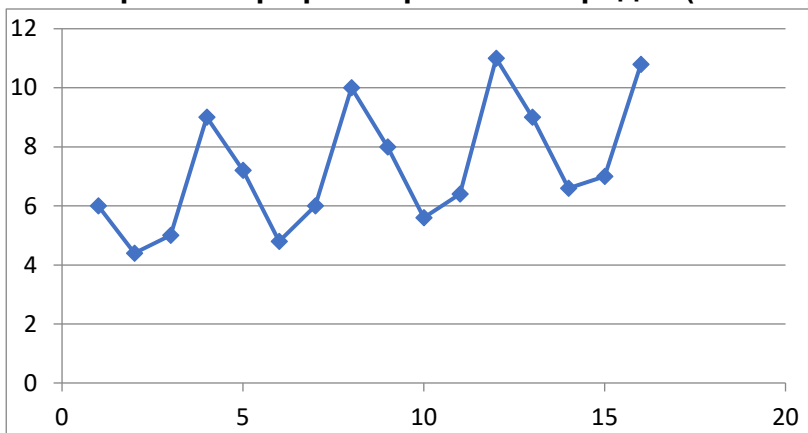
Задача. Имеются поквартальные данные об объеме потребления электроэнергии жителями региона за последние четыре года, млн. кВт·ч:

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
y_t	6	4,4	5	9	7,2	4,8	6	10	8	5,6	6,4	11	9	6,6	7	10,8

1. Построить график временного ряда.
2. Определить структуру временного ряда, используя автокорреляционную функцию.
3. Построить аддитивную модель временного ряда.
4. Определить качество построенной модели.
5. Сделать прогноз потребления электроэнергии на первое полугодие ближайшего следующего года.

Решение

1. Построение графика временного ряда (в Excel)

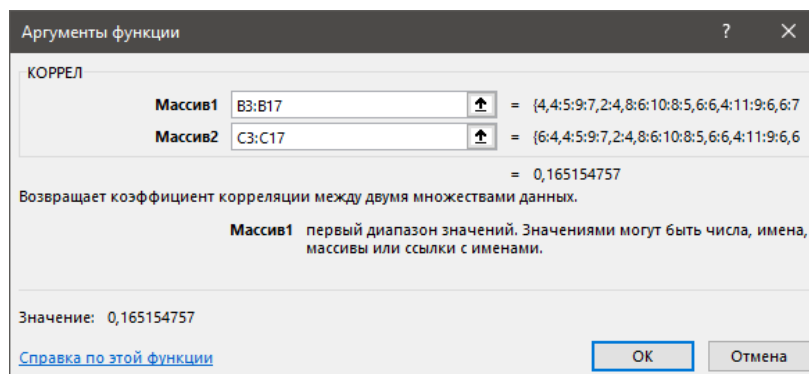


Вывод: Графический анализ временного ряда показывает наличие трендовой компоненты (имеется устойчивая тенденция роста показателя). Также временной ряд содержит сезонные колебания периодичностью 4. Значение изучаемого показателя в I и IV кварталах выше, чем во II и III кварталах. По графику можно установить наличие приблизительно равной амплитуды колебаний. Это свидетельствует о соответствии ряда аддитивной модели.

2. Определение структуры временного ряда с помощью автокорреляционной функции

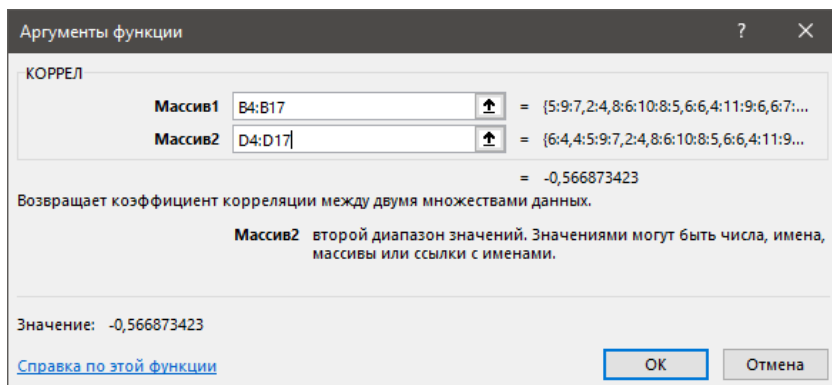
Рассчитаем коэффициенты автокорреляции 1, 2, ..., 8 порядков, используя статистическую функцию Коррел (Excel). Результаты вычислений представим в таблице 1:

а) Для вычисления коэффициента автокорреляции первого порядка $r(1)$ добавляем столбец y_{t-1} . По 15 наблюдениям выполняем расчет, используя функцию Коррел:



Полученное значение $r(1)=0,165$ говорит о слабой зависимости текущих и предыдущих уровней ряда.

б) Рассчитаем коэффициент автокорреляции второго порядка $r(2)$ по 14 наблюдениям, используя ряды y_t и y_{t-2} .



в) Продолжив расчеты аналогичным образом, получим автокорреляционную функцию этого ряда (см таблицу 1).

Таблица 1 – Расчет 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}
1	6								
2	4,4	6							
3	5	4,4	6						
4	9	5	4,4	6					
5	7,2	9	5	4,4	6				
6	4,8	7,2	9	5	4,4	6			
7	6	4,8	7,2	9	5	4,4	6		
8	10	6	4,8	7,2	9	5	4,4	6	
9	8	10	6	4,8	7,2	9	5	4,4	6
10	5,6	8	10	6	4,8	7,2	9	5	4,4
11	6,4	5,6	8	10	6	4,8	7,2	9	5
12	11	6,4	5,6	8	10	6	4,8	7,2	9
13	9	11	6,4	5,6	8	10	6	4,8	7,2
14	6,6	9	11	6,4	5,6	8	10	6	4,8
15	7	6,6	9	11	6,4	5,6	8	10	6
16	10,8	7	6,6	9	11	6,4	5,6	8	10
Порядок(лаг)		1	2	3	4	5	6	7	8
Коэффициент автокорреляции $r(\tau)$		0,165	-0,567	0,114	0,983	0,119	-0,722	-0,003	0,974

Вывод: наиболее высокими оказались коэффициенты автокорреляции уровней 4, 8 порядков, т.е. временной ряд содержит циклические колебания в 4 периода.

3. Построение аддитивной модели временного ряда.

Общий вид аддитивной модели: $Y = T + S + E$, т.е. каждый уровень временного ряда представлен как сумма трендовой T , сезонной S и случайной E компонент.

Аддитивная модель применяется в том случае, когда амплитуду колебаний уровней ряда относительно тренда можно считать постоянной.

Построение аддитивной модели сводится к расчету значений T , S и E для каждого уровня ряда.

Процесс построения модели включает следующие шаги:

1. Выравнивание исходного ряда методом скользящей средней.
2. Расчет значений сезонной компоненты S .
3. Устранение сезонной компоненты из исходных уровней ряда и получение выровненных данных $(T + E)$.
4. Аналитическое выравнивание уровней $(T + E)$ и расчет значений T с использованием уравнения тренда.
5. Расчет полученных по модели значений $(T + S)$.
6. Расчет абсолютных и/или относительных ошибок.

Данный временной ряд имеет сезонную составляющую и амплитуда колебаний приблизительно одинаковая от цикла к циклу. Следовательно, будем строить аддитивную модель.

Расчет оценок сезонной составляющей в аддитивной модели приведены в таблице 2.

Таблица 2

№ квартала	Потребители электроэнергии y_t	Итого за квартал	Скользящая средняя за 4 квартала	Центрированная скользящая средняя	Оценка сезонной компоненты
1	2	3	4	5	6
1	6	—	—	—	—
2	4,4	24,4	6,10	—	—
3	5	25,6	6,40	6,25	—1,250
4	9	26	6,50	6,45	2,550
5	7,2	27	6,75	6,625	0,575
6	4,8	28	7,00	6,875	—2,075
7	6	28,8	7,20	7,1	—1,100
8	10	29,6	7,40	7,3	2,700
9	8	30	7,50	7,45	0,550
10	5,6	31	7,75	7,625	—2,025
11	6,4	32	8,00	7,875	—1,475
12	11	33	8,25	8,125	2,875
13	9	33,6	8,40	8,325	0,675
14	6,6	33,4	8,35	8,375	—1,775
15	7	—	—	—	—
16	10,8	—	—	—	—

Поясним, как заполняется эта таблица:

– столбец 3 – суммируем расход электроэнергии за 4 квартала, т.е.

$$6 + 4,4 + 5 + 9 = 6,10,$$

$$4,4 + 5 + 9 + 7,2 = 25,6 \text{ и т.д.};$$

– столбец 4 – находим скользящую среднюю за 4 квартала:

$$24,4 : 4 = 6,1,$$

$$25,6 : 4 = 6,40 \text{ и т.д.};$$

– столбец 5 – центрированная скользящая средняя:

$$(6,10 + 6,40) : 2 = 6,25,$$

$$(6,40 + 6,50) : 2 = 6,45 \text{ и т.д.};$$

– столбец 6 – оценка сезонной компоненты, начинаем с $t = 3$ – из значений столбца 2 вычитаем значения столбца 5:

$$5 - 6,25 = -1,250,$$

$$9 - 6,45 = 2,550 \text{ и т.д.}$$

Скорректируем сезонную составляющую.

Возьмем оценку из предыдущей табл. и заполним следующую таблицу.

Таблица 3

Квар- талы	Годы				Итого за квар- тал	Средняя оценка сезонной состав- ляющей $\bar{S}_i$	Скорректиро- ванная сезон- ная состав- ляющая $S_i = \bar{S}_i - k$
	1	2	3	4			
1	–	0,575	0,550	0,675	1,800	0,600	0,581
2	–	–2,075	–2,025	–1,777	–5,875	–1,985	–1,977
3	–1,250	–1,100	–1,475	–	–3,825	–1,275	–1,294
4	2,550	2,700	2,875	–	8,125	2,708	2,690
Σ						0,075	0

В моделях с сезонной компонентой обычно предполагается, что сезонные воздействия за период взаимопогашаются, т.е. $\sum_{i=1}^4 \bar{S}_i = 0$.

В нашем случае сумма не равна нулю ($\sum_{i=1}^4 \bar{S}_i = 0,075$). Тогда корректирующий коэффициент будет равен:

$$k = \frac{\sum_{i=1}^4 \bar{S}_i}{4} = \frac{0,075}{4} = 0,01875.$$

Для дальнейших расчетов оставим 3 знака после запятой:

$$\begin{aligned} S_1 &= \bar{S}_1 - k = 0,600 - 0,0185 = 0,581, \\ S_2 &= \bar{S}_2 - k = -1,985 - 0,0185 = -1,977, \\ S_3 &= \bar{S}_3 - k = -1,275 - 0,0185 = -1,294, \\ S_4 &= \bar{S}_4 - k = 2,708 - 0,0185 = 2,690. \end{aligned}$$

Теперь элиминируем (исключим) влияние сезонной составляющей. Расчет выровненных значений T и ошибок E в аддитивной модели приведен в табл. 4.

Исключим сезонную составляющую S из модели:

$$Y = T + S + E, \text{ следовательно } Y - S = T + E.$$

Полученный ряд (столбец 5) содержит уже только тенденцию и случайную составляющую.

Для данных $(t, T + E)$ построим тренд линейного типа:

$$T = 5,7155 + 0,1864t.$$

Теперь по этой модели найдем значения T для каждого момента времени t (столбец 6).

Далее для каждого уровня t найдем значения $T + S$ (значения в столбце 7 равны сумме данных 6-ого и 4-ого столбцов).

Рассчитаем абсолютную ошибку по полученной аддитивной модели:

$$E = Y - (T + S_t)$$

и найдем ее квадрат.

Таблица 4.

№	$t_{\text{кв}}$	y_t	s_t	$T + E = y_t - s_t$	$T = 5,716 + 0,186t$	$T + s_t$	$E = y_t - (T + s_t)$	E^2	$(y_t - \bar{y}_t)^2$
1	2	3	4	5	6	7	8	9	10
1	1	6,0	0,581	5,419	5,902	6,483	-0,483	0,2333	1,69
	2	4,4	-1,977	6,337	6,088	4,111	0,289	0,0835	8,41
	3	5,0	-1,294	6,294	6,275	4,981	0,019	0,0004	5,29
	4	9,0	2,690	6,310	6,461	9,151	-0,151	0,0228	2,89
2	5	7,2	0,581	6,619	6,648	7,229	-0,029	0,0008	0,01
	6	4,8	-1,977	6,777	6,834	4,857	-0,057	0,0032	6,25
	7	6,0	-1,294	7,294	7,020	5,726	0,274	0,0751	1,69
	8	10,0	2,690	7,310	7,207	9,897	0,103	0,0106	7,29
3	9	8,0	0,581	7,419	7,393	7,974	0,026	0,0007	0,49
	10	5,6	-1,977	7,577	7,580	5,603	-0,002	0,0000	2,89
	11	6,4	-1,294	7,694	7,766	6,472	-0,072	0,0052	0,81
	12	11,0	2,690	8,310	7,952	10,642	0,358	0,1282	13,69
4	13	9,0	0,581	8,419	8,139	8,720	0,280	0,0784	2,89
	14	6,6	-1,977	8,577	8,325	6,348	0,252	0,0635	0,49
	15	7,0	-1,294	8,294	8,512	7,218	-0,217	0,0471	0,09
	16	10,8	2,690	8,110	8,698	11,388	-0,588	0,3456	12,25
Σ		116,8						1,0984	67,12
$\bar{y}_t$		7,3							

4. Определение качества построенной модели.

По отношению к общей сумме квадратов отклонений уровней ряда от его среднего уровня $\bar{y}_t = 7,2375$ составит:

$$\frac{E^2}{\sum (y_t - \bar{y})^2} = \frac{1,0984}{67,12} \cdot 100 = 0,0164 \cdot 100 = 1,64\%.$$

Тогда объясненная вариация

$$\left(1 - \frac{1,0984}{67,12}\right) \cdot 100 = 98,36\%.$$

Вывод: полученная аддитивная модель объясняет 98,36% общей вариации уровней временного ряда потребления электроэнергии за шестнадцать кварталов. Эту модель логично использовать для прогноза.

5. Прогноз потребления электроэнергии на первое полугодие следующего года.

Предположим, надо спрогнозировать потребление электроэнергии на первое полугодие ближайшего следующего года, т.е. для $t = 17$ и $t = 18$. Для прогноза используем полученную модель

$$T = 5,7155 + 0,1864 \cdot t.$$

Подставим значения $t = 17$ и $t = 18$ в модель:

$$T_{17} = 5,7155 + 0,1864 \cdot 17 = 8,884,$$

$$T_{18} = 5,7155 + 0,1864 \cdot 18 = 9,071.$$

Значение сезонной составляющей для 1-го квартала $S_1 = 0,581$, для 2-го квартала $S_2 = -1,977$.

Прогноз с учетом сезонной составляющей (т.е. тренд + сезонная составляющая):

$$F_{17} = T_{17} + 0,581 = 8,884 + 0,561 = 9,465,$$

$$F_{18} = T_{18} - 1,977 = 9,071 - 1,977 = 7,094.$$

Следовательно, на следующее полугодие надо запланировать 16,559 млн. кВт/ч.