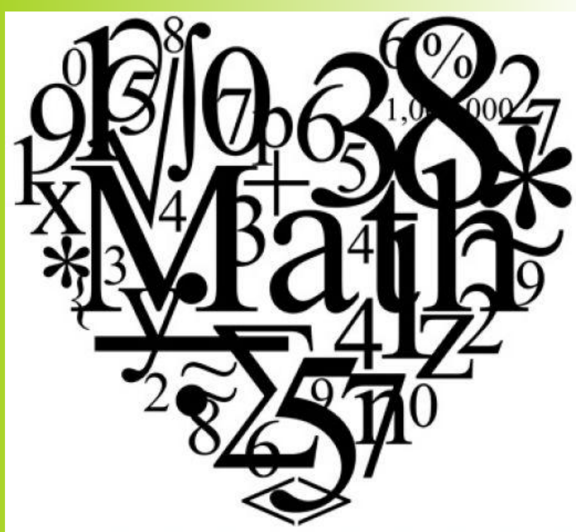
Лекция №7

Тема 1. Модели временных рядов



ПЛАН ЛЕКЦИИ

1. Понятие временного ряда.
Общий вид модели временного ряда.
2. Проверка гипотезы существования тенденции.
3. Стационарные временные ряды и их характеристики.
4. Выявление структуры временного ряда.

1. ПОНЯТИЕ ВРЕМЕННОГО РЯДА. ОБЩИЙ ВИД МОДЕЛИ ВРЕМЕННОГО РЯДА.

Временной (динамический) ряд

Это последовательность наблюдений Y некоторого признака (случайной величины) в последовательные моменты времени.

1. ПОНЯТИЕ ВРЕМЕННОГО РЯДА. ОБЩИЙ ВИД МОДЕЛИ ВРЕМЕННОГО РЯДА.

Отдельные наблюдения называются уровнями ряда.

Обозначим их

$$y_t \ (t=1,2,\ldots,n), \qquad (1)$$

где n – число уровней

Каждый уровень временного ряда $y_t = y(t)$ формируется под влиянием большого числа факторов, условно разделяемых на три группы

- Длительные (формируют тенденцию ряда)
- Кратковременные – циклические колебания
- Случайные факторы

1. ПОНЯТИЕ ВРЕМЕННОГО РЯДА. ОБЩИЙ ВИД МОДЕЛИ ВРЕМЕННОГО РЯДА.

Факторы, которые действуют в течение длительного времени, оказывают определяющее влияние и формируют основную тенденцию временного ряда – **тренд T(t)**. Это **трендовая компонента ряда**.

Периодические факторы формируют циклические колебания ряда – **сезонную (циклическую) компоненту S(t)**. Например, колебания цен на сельхозпродукцию носят сезонный характер.

Случайные факторы отражаются случайными изменениями уровней ряда и образуют **случайную компоненту ε(t).**

В большинстве случаев фактический уровень временного ряда можно представить как сумму или произведение трендовой, циклической и случайной компонент.

1. ПОНЯТИЕ ВРЕМЕННОГО РЯДА. ОБЩИЙ ВИД МОДЕЛИ ВРЕМЕННОГО РЯДА.

Модель, в которой временной ряд представлен суммой этих компонент:

$$y(t) = T(t)+S(t)+\varepsilon(t) \quad (2)$$

или

$$Y_t=T_t+S_t+\varepsilon_t \quad (3)$$

называется ***аддитивной моделью временного ряда***

Модель, в которой временной ряд представлен произведением этих компонент:

$$y(t) = T(t)*S(t)*\varepsilon(t) \quad (4)$$

или

$$Y_t=T_t*S_t*\varepsilon_t \quad (5)$$

называется ***мультипликативной моделью временного ряда***

Выбор типа модели обычно регламентируется поведением амплитуды сезонных колебаний

амплитуда сезонных колебаний близка к постоянной → **Аддитивная модель**

амплитуда сезонных колебаний убывает или возрастает → **Мультипликативная модель**

1. ПОНЯТИЕ ВРЕМЕННОГО РЯДА. ОБЩИЙ ВИД МОДЕЛИ ВРЕМЕННОГО РЯДА.

Задача эконометрического исследования временного ряда

↓

Определение количественного выражения каждой из перечисленных компонент с тем, чтобы можно было полученную модель использовать для прогнозирования будущих значений ряда

1. ПОНЯТИЕ ВРЕМЕННОГО РЯДА. ОБЩИЙ ВИД МОДЕЛИ ВРЕМЕННОГО РЯДА.

- Для моделирования временных рядов используют модели парной линейной и нелинейной регрессии, множественной линейной и нелинейной регрессии и специальные модели.
- ***Для построения трендов используют различные функции, основные:***

Линейный тренд	$\hat{y}_t = a + bt$	**(6)**
Гипербола	$\hat{y}_t = a + \frac{b}{t}$	**(7)**
Экспоненциальный тренд	$\hat{y}_t = e^{a+bt}$ или $\hat{y}_t = a * b^t$	**(8)**
Степенной тренд	$\hat{y}_t = a * t^b$	**(9)**
Многочлен (второго и более порядков)	$\hat{y}_t = a + b_1 t + b_2 t^2 + \ldots + b_k t^k$	**(10)**

Этапы анализа временных рядов:

графическое представление и описание поведения временного ряда

выделение и удаление закономерных (неслучайных) составляющих временного ряда (тренда, сезонных и циклических составляющих)

сглаживание и фильтрация (удаление низко- или высокочастотных составляющих временного ряда)

исследование случайной составляющей временного ряда, построение и проверка адекватности математической модели для ее описания

прогнозирование развития изучаемого процесса на основе имеющегося временного ряда

исследование взаимосвязи между различными временными рядами

2. ПРОВЕРКА ГИПОТЕЗЫ СУЩЕСТВОВАНИЯ ТЕНДЕНЦИИ.

Прогнозирование временных рядов целесообразно начинать с построения графика исследуемого показателя. Однако в нем не всегда прослеживается присутствие тренда.

Поэтому в этих случаях необходимо выяснить, существует ли тенденция во временном ряду или она отсутствует.

2. ПРОВЕРКА ГИПОТЕЗЫ СУЩЕСТВОВАНИЯ ТЕНДЕНЦИИ.

Для временного ряда рассмотрим критерий «восходящих и нисходящих» серий, согласно которому тенденция определяется по следующему алгоритму:

1

Для исследуемого временного ряда $y_1, y_2, \dots, y_t, \dots, y_n$ определяется последовательность знаков, исходя из условий:

$$\delta_i = \begin{cases} +, & \text{если} \quad y_{t+1} -_t > 0. \\ -, & \text{если} \quad y_{t+1} - y_t < 0. \end{cases} \qquad (11)$$

При этом, если последующее наблюдение равно предыдущему, то учитывается только одно наблюдение.

2

Подсчитывается число серий *v* (*n*). Под <u>*серией*</u> понимается последовательность подряд расположенных плюсов или минусов, причем один плюс или один минус считается серией.

2. ПРОВЕРКА ГИПОТЕЗЫ СУЩЕСТВОВАНИЯ ТЕНДЕНЦИИ.

3 **Определяется протяженность самой длинной серии $l_{max}(n)$.**

4 **По таблице, приведенной ниже, находится значение $l(n)$.**

ТАБЛИЦА 1

Длина ряда *n*	$n \leq 26$	$26 < n < 153$	$153 < n < 170$
Значение $l(n)$	**5**	**6**	**7**

5 **Если нарушается хотя бы одно из следующих неравенств, то гипотеза об отсутствии тренда отвергается с доверительной вероятностью 0,95:**

$$\begin{cases} \upsilon(n) > \left[1/3 \cdot (2n-1) - 1{,}96 \cdot \sqrt{(16n-29)/90}\right], \\ l_{max}(n) \leq l(n). \end{cases} \qquad (12)$$

Квадратные скобки неравенства в (12) означают целую часть числа.

2. ПРОВЕРКА ГИПОТЕЗЫ СУЩЕСТВОВАНИЯ ТЕНДЕНЦИИ.

Пример 1

Дана динамика ежеквартального выпуска продукции фирмы в ден. ед. С помощью критерия «восходящих и нисходящих» серий сделать вывод о присутствии или отсутствии тренда. Доверительную вероятность принять равной 0,95.

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
y_t	10	14	7	16	15	17	16	20	17	7	15	16	20	14	19	21

Решение.

Определим последовательность знаков.

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
y_t	10	14	7	16	15	17	16	20	17	7	15	16	20	14	19	21
δ_t		+	–	+	–	+	–	+	–	–	+	+	+	–	+	+

Число серий $\upsilon(n) = 11$, протяженность самой длинной серии $l_{max}(n) = 3$, по таблице $l(n) = 5$.

2. ПРОВЕРКА ГИПОТЕЗЫ СУЩЕСТВОВАНИЯ ТЕНДЕНЦИИ.

Запишем систему неравенств:

$$\begin{cases} 11 > \left[1/3 \cdot (2 \cdot 16 - 1) - 1{,}96 \cdot \sqrt{(16 \cdot 16 - 29)/90}\right], \\ 3 \le 5, \end{cases}$$

$$\begin{cases} 11 > 7, \\ 3 \le 5. \end{cases}$$

Оба неравенства выполняются, поэтому тренд в динамике выпуска продукции фирмы отсутствует с доверительной вероятностью 0,95.

3. СТАЦИОНАРНЫЕ ВРЕМЕННЫЕ РЯДЫ И ИХ ХАРАКТЕРИСТИКИ.

Важное значение в анализе временных рядов имеют стационарные временные ряды, вероятностные свойства которых не изменяются во времени.

	Временной ряд $y_t(t=1,2,\ldots,n)$ называется *стационарным*, если совместное распределение вероятностей *n* наблюдений $y_1,y_2,\ldots,y_n$ такое же, как и *n* наблюдений $y_{1+\tau},y_{2+\tau},\ldots,y_{n+\tau}$ при любых n, t и τ. Другими словами, свойства строго стационарных рядов y_t не зависят от момента t, т.е. закон распределения и его числовые характеристики не зависят от t.

3. СТАЦИОНАРНЫЕ ВРЕМЕННЫЕ РЯДЫ И ИХ ХАРАКТЕРИСТИКИ.

Следовательно, математическое ожидание $M_y(t)=a$, среднее квадратическое отклонение $\sigma_y(t)=\sigma$ могут быть оценены по наблюдениям с помощью формул:

$$\bar{y}_t = \frac{\sum_{t=1}^{n} y_t}{n}, \tag{13}$$

$$s_t^2 = \frac{\sum_{t=1}^{n} (y_t - \bar{y}_t)^2}{n} \tag{14}$$

3. СТАЦИОНАРНЫЕ ВРЕМЕННЫЕ РЯДЫ И ИХ ХАРАКТЕРИСТИКИ.

Между значениями временного ряда на отдельных его участках может существовать корреляционная связь

Корреляционная зависимость между последовательными уровнями временного ряда называется <u>автокорреляцией уровней временного ряда</u>

3. СТАЦИОНАРНЫЕ ВРЕМЕННЫЕ РЯДЫ И ИХ ХАРАКТЕРИСТИКИ.

Степень тесноты связи между последовательностями наблюдений временного ряда $y_1, y_2, \ldots, y_n$ и $y_{1+r}, y_{2+r}, \ldots, y_{n+r}$ (сдвинутых относительно друг друга на τ единиц, или, как говорят, с лагом τ) может быть определена с помощью коэффициента корреляции:

$$p(\tau) = \frac{M\left[\left(y_t - a\right)\left(y_{1+r} - a\right)\right]}{\sigma^2} \qquad (13)$$

Т.к. $p(\tau)$ измеряет корреляцию между членами одного и того же ряда, его называют *коэффициентом автокорреляции*, а зависимость $p(\tau)$ – *автокорреляционной функцией*.

3. СТАЦИОНАРНЫЕ ВРЕМЕННЫЕ РЯДЫ И ИХ ХАРАКТЕРИСТИКИ.

Статистической оценкой **р(τ)** является **выборочный коэффициент автокорреляции r(τ)**, определяемый по формуле:

$$r(\tau)=\frac{(n-\tau)\sum_{t=1}^{n-r} y_t y_{t+r}-\sum_{t=1}^{n-r} y_t \sum_{t=1}^{n-r} y_{t+r}}{\sqrt{(n-\tau)\sum_{t=1}^{n-r} y_t^2-\left(\sum_{t=1}^{n-r} y_t\right)^2}\sqrt{(n-\tau)\sum_{t=1}^{n-r} y_{t+r}^2-\left(\sum_{t=1}^{n-r} y_{t+r}\right)^2}} \qquad \textbf{(15)}$$

3. СТАЦИОНАРНЫЕ ВРЕМЕННЫЕ РЯДЫ И ИХ ХАРАКТЕРИСТИКИ.

Число периодов **τ**, для которых рассчитывается коэффициент автокорреляции, называют **лагом**

Функцию **r(τ)** называют **выборочной автокорреляционной функцией**

График функции **r(τ)** называют **коррелограммой**. Она позволяет определить лаг, при котором связь между уровнями является наиболее тесной.

При расчете **r(τ)** следует помнить, что с увеличением **τ** число **n-τ** пар наблюдений y_t, y_{t+r} уменьшается, поэтому лаг **τ** должен быть таким, чтобы число **n-τ** было достаточным для определения **r(τ)**. Обычно ориентируются на соотношение

$$\tau \le n/4 \cdot$$

3. СТАЦИОНАРНЫЕ ВРЕМЕННЫЕ РЯДЫ И ИХ ХАРАКТЕРИСТИКИ.

Наряду с автокорреляционной функцией при исследовании стационарных рядов рассматривается **частная автокорреляционная функция** $\rho_{част}(\tau)$, где $\rho_{част}(\tau)$ есть **частный коэффициент корреляции между членами временного ряда** y_t, y_{t+r} , т.е. коэффициент корреляции между y_t, y_{t+r} и при устранении влияния промежуточных между y_t и y_{t+r} членов.

4. ВЫЯВЛЕНИЕ СТРУКТУРЫ ВРЕМЕННОГО РЯДА.

Автокорреляционную функцию обычно используют для выявления наличия или отсутствия трендовой и сезонной компонент:

если наиболее высоким оказывается коэффициент автокорреляции 1-го порядка (r_1), то исследуемый ряд содержит только тенденцию – трендовую компоненту	если наиболее высоким оказался коэффициент r_k (при $k>1$), то ряд содержит также сезонную компоненту (циклические колебания) с периодом в k моментов времени	если ни один из коэффициентов автокорреляции не оказался значимым, можно предположить: либо ряд не содержит тенденции и циклических колебаний, либо ряд содержит сильную нелинейную тенденцию, для выявления которой требуется дополнительный анализ

Пример 1

Пусть имеются данные об объемах потребления электроэнергии жителями региона за 16 кварталов (первые два столбца таблицы 2). Определить структуру данного временного ряда.

ТАБЛИЦА 2

t	yt	yt-1	yt-2	yt-3	yt-4
1	6,0	-	-	-	-
2	4,4	6,0	-	-	-
3	5,0	4,4	6,0	-	-
4	9,0	5,0	4,4	6,0	-
5	7,2	9,0	5,0	4,4	6,0
6	4,8	7,2	9,0	5,0	4,4
7	6,0	4,8	7,2	9,0	5,0
8	10,0	6,0	4,8	7,2	9,0
9	8,0	10,0	6,0	4,8	7,2
10	5,6	8,0	10,0	6,0	4,8
11	6,4	5,6	8,0	10,0	6,0
12	11,0	6,4	5,6	8,0	10,0
13	9,0	11,0	6,4	5,6	8,0
14	6,6	9,0	11,0	6,4	5,6
15	7,0	6,6	9,0	11,0	6,4
16	10,8	7,0	6,6	9,0	11,0

4. ВЫЯВЛЕНИЕ СТРУКТУРЫ ВРЕМЕННОГО РЯДА.

Построим график для y_t

РИСУНОК 1

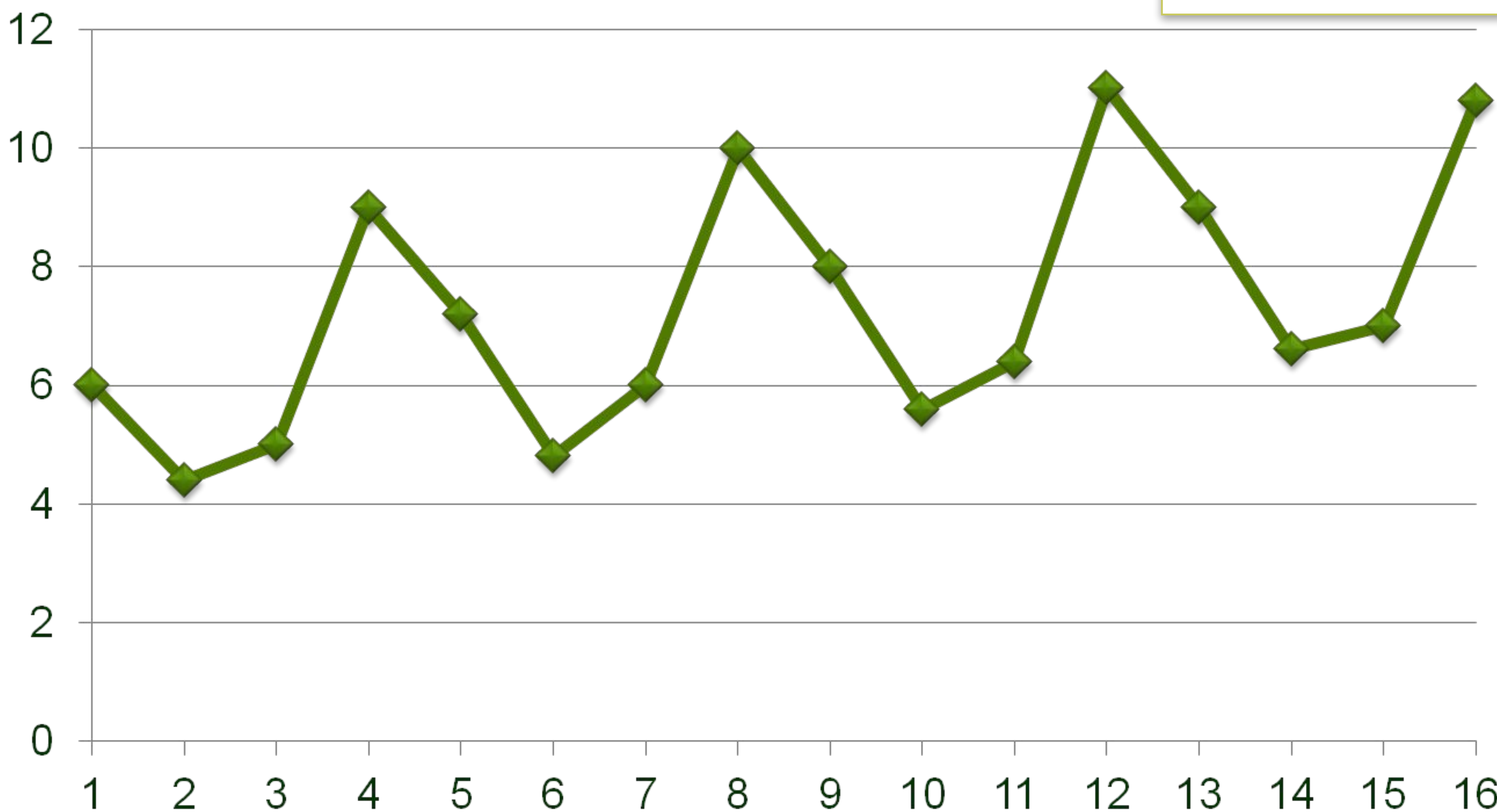


Рис.1. Потребление электроэнергии жителями региона

4. ВЫЯВЛЕНИЕ СТРУКТУРЫ ВРЕМЕННОГО РЯДА.

По формуле для расчета линейного коэффициента корреляции определим коэффициент автокорреляции первого порядка r_1 (столбцы y_t и y_{t-1})

Расчет этого коэффициента проводится по 15 наблюдениям, а не по 16 (т.к. из-за временного сдвига столбец y_{t-1} содержит 15 значений)

Получим значение $r_1=0,165$

Это значение говорит о слабой зависимости текущих уровней ряда от непосредственно предшествующих им уровней

Структура ряда такова, что каждый следующий уровень y_t зависит от уровней y_{t-4} и y_{t-2} в гораздо большей степени, чем от уровня y_{t-1}

4. ВЫЯВЛЕНИЕ СТРУКТУРЫ ВРЕМЕННОГО РЯДА.

Рассчитаем остальные коэффициенты автокорреляции (второго, третьего и далее до восьмого порядка включительно). Получим автокорреляционную функцию этого ряда

ТАБЛИЦА 3

Лаг	Коэффициент автокорреляции уровней	Коррелограмма
1	0,165154	**
2	0,566873	*******
3	0,113558	*
4	0,983025	************
5	0,118711	*
6	0,722046	*********
7	0,003367	
8	0,973848	************

Анализ значений этой функции приводит к выводу о существовании, помимо линейной тенденции, сезонных колебаний с периодом в 4 квартала. Этот вывод подтверждается и графическим анализом структуры ряда (рис.1)