Тема 5. Анализ временных рядов

1. Понятие временного ряда и его основные компоненты.

2. Построение аддитивной модели.

3. Построение мультипликативной модели.
4. Моделирование тенденции временного ряда при наличии структурных изменений

1 вопрос

Временной ряд - это совокупность значений какого-либо показателя за несколько последовательных моментов (периодов) времени (y_t).

Модели, построенные по временным рядам, называются ***моделями временных рядов***. Параметры таких моделей оцениваются специальными методами, разработанными на основе традиционных методов регрессионного анализа.

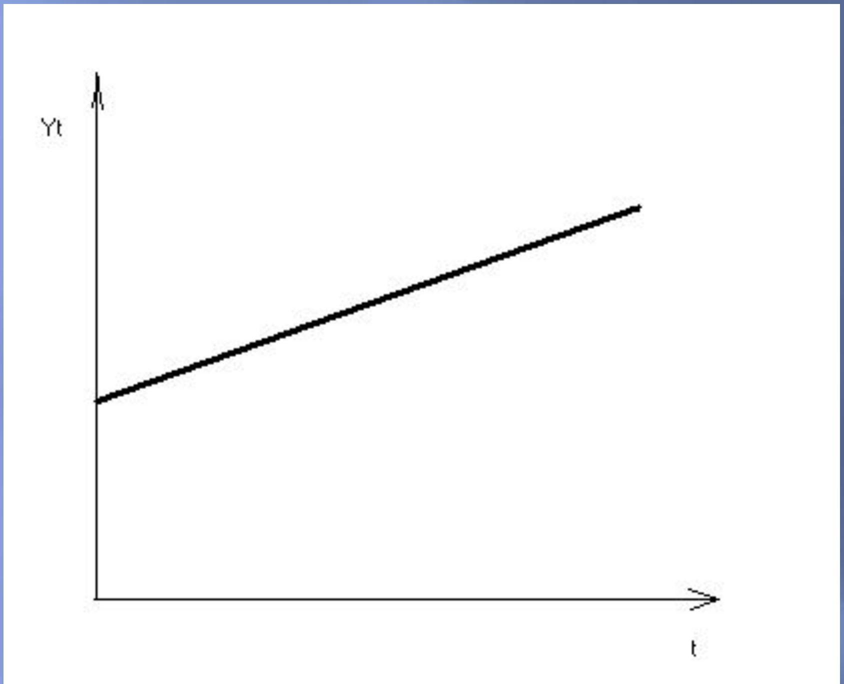
В чем особенность временного ряда?

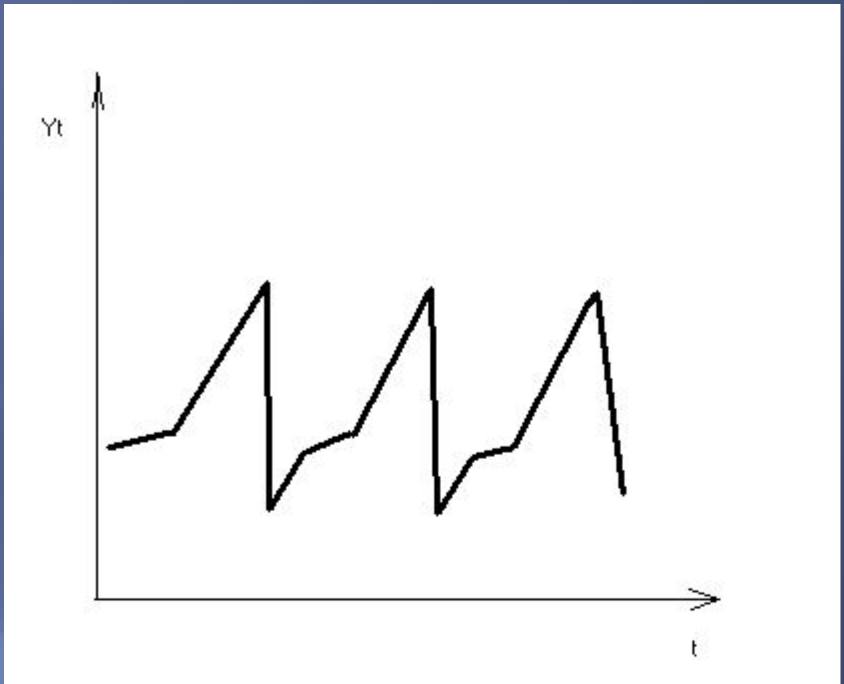
Каждый уровень(Y_t) формируется под влиянием трех групп факторов

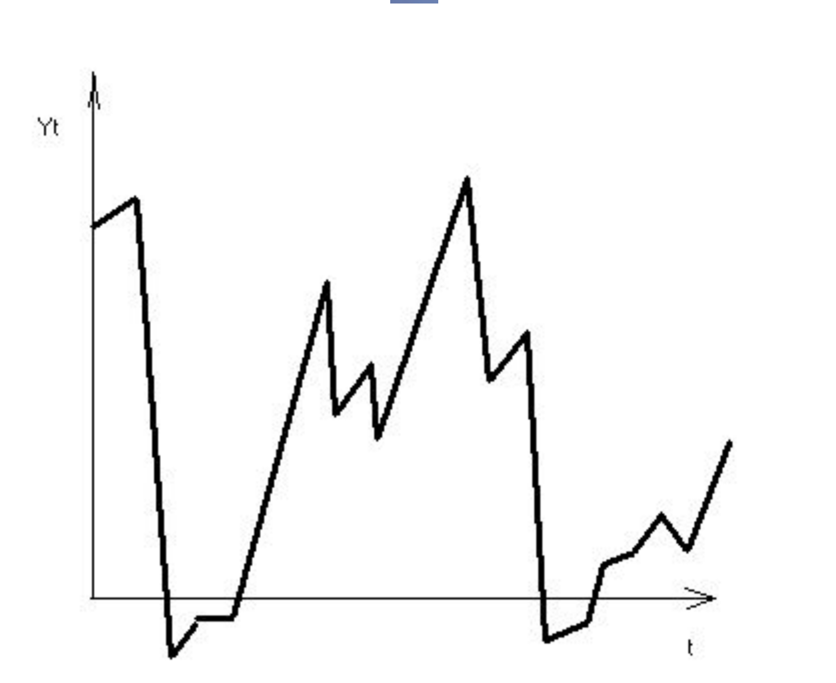
Факторы, формирующие тенденцию (T_t)

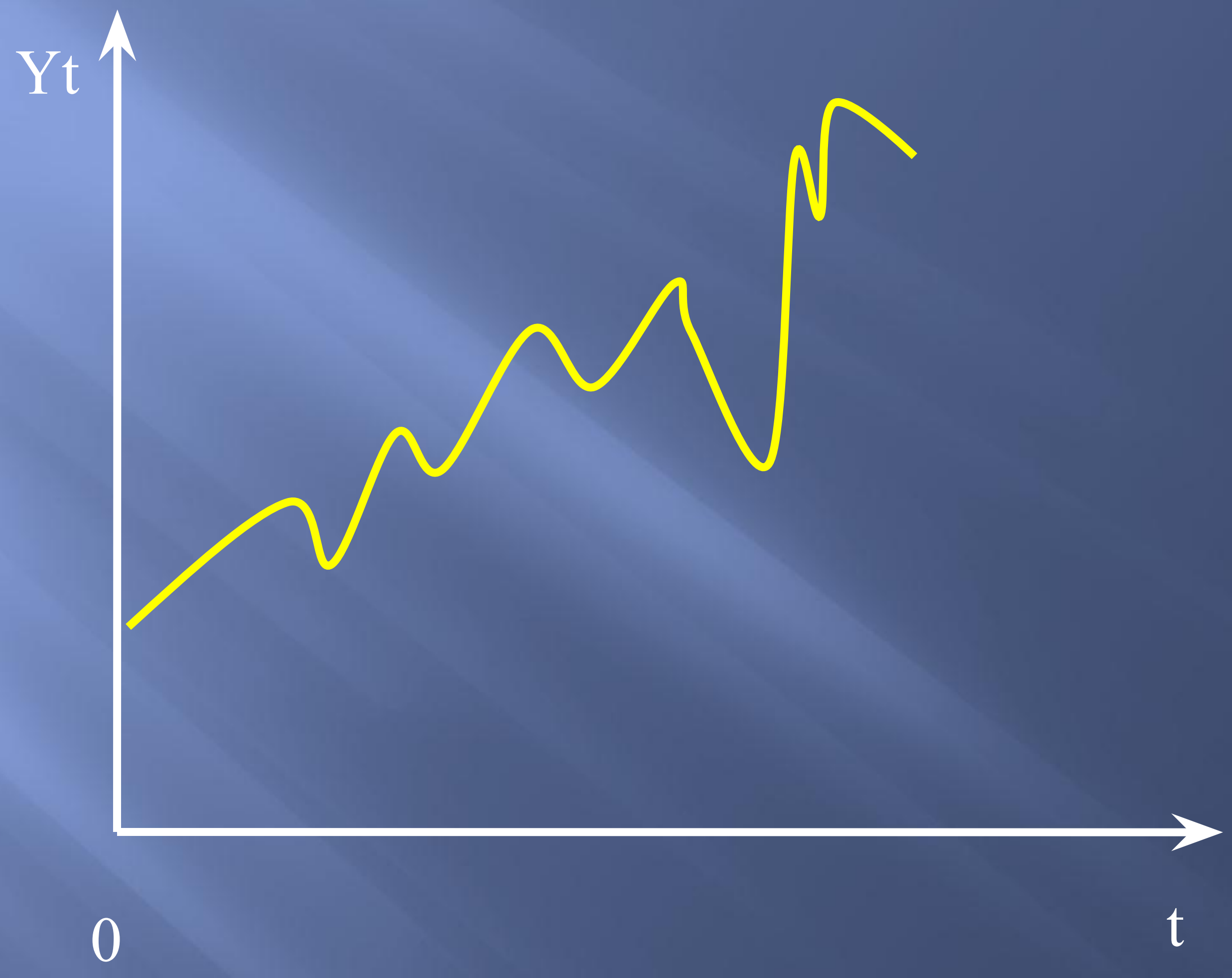
Факторы, формирующие циклические колебания (S_t)

Случайные факторы (ε_t)


Yt
t


Yt
t


Yt
t



Реальные данные чаще всего содержат все три компоненты

Модели временного ряда

- Модель, в которой временной ряд представлен как сумма его компонент, называется аддитивной моделью временного ряда:

$$y_t = T_t + S_t + e_t$$

- Модель, в которой временной ряд представлен как произведение его компонент, называется мультипликативной моделью временного ряда:

$$y_t = T_t \cdot S_t \cdot e_t$$

Основная задача эконометрического исследования временного ряда – выявление и количественное измерение тенденции, циклической и случайной компонент, с тем, чтобы использовать информацию для получения прогнозных оценок или при построении моделей взаимосвязи двух или более временных рядов.

При наличии тенденции и циклических колебаний значения каждого последующего уровня ряда зависят от предыдущих значений

Автокорреляция уровней ряда

- ✔Корреляционную зависимость между последовательными уровнями временного ряда называют ***автокорреляцией уровней ряда.***

- ✔Число периодов, по которым рассчитывается коэффициент автокорреляции, называется ***лагом.*** Максимальный лаг должен быть ***не больше n/4.***

- ✔Последовательность коэффициентов автокорреляции уровней первого, второго и т. д. порядков называют ***автокорреляционной функцией временного ряда.***

- ✔График зависимости ее значений от величины лага называется ***коррелограммой.***

Коэффициент автокорреляции уровней ряда первого порядка

$$r_1 = \frac{\sum_{t=2}^{n} (y_t - \bar{y}_1) \cdot (y_{t-1} - \bar{y}_2)}{\sqrt{\sum_{t=2}^{n} (y_t - \bar{y}_1)^2 \cdot \sum_{t=2}^{n} (y_{t-1} - \bar{y}_2)^2}}$$

$$\bar{y}_1 = \frac{\sum_{t=2}^{n} y_t}{n-1}; \bar{y}_2 = \frac{\sum_{t=2}^{n} y_{t-1}}{n-1}$$

Коэффициент автокорреляции уровней ряда второго порядка

$$r_2 = \frac{\sum_{t=3}^{n} (y_t - \bar{y}_3) \cdot (y_{t-2} - \bar{y}_4)}{\sqrt{\sum_{t=3}^{n} (y_t - \bar{y}_3)^2 \cdot \sum_{t=3}^{n} (y_{t-2} - \bar{y}_4)^2}}$$

$$\bar{y}_3 = \frac{\sum_{t=3}^{n} y_t}{n-2}; \bar{y}_4 = \frac{\sum_{t=3}^{n} y_{t-2}}{n-2}$$

Свойства коэффициента автокорреляции

- Характеризует тесноту только линейной связи текущего и предыдущего уровней ряда.

- По знаку коэффициента нельзя делать вывод о возрастающей или убывающей тенденции в уровнях ряда.

Потребление электроэнергии жителями региона, млн. кВт*ч					
t	Yt	Yt-1	Yt-2	Yt-3	Yt-4
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
7	6	4,8	7,2	9	5
8	10	6	4,8	7,2	9
9	8	10	6	4,8	7,2
10	5,6	8	10	6	4,8
11	6,4	5,6	8	10	6
12	11	6,4	5,6	8	10
13	9	11	6,4	5,6	8
14	6,6	9	11	6,4	5,6
15	7	6,6	9	11	6,4
16	10,8	7	6,6	9	11

Коррелограмма временного ряда потребления электроэнергии		
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	************

При помощи анализа автокорреляционной функции и коррелограммы можно выявить структуру ряда

Выводы о структуре временного ряда

- **Самый высокий r_1**
  - Ряд содержит только тенденцию
- **Самый высокий r_T**
  - Ряд содержит циклические колебания с периодичностью в T моментов времени
- **Ни один r не значим**
  - Ряд не содержит тенденцию и циклические колебания, либо имеет сильную нелинейную тенденцию

Методы выявления основной тенденции временного ряда

- Сглаживание или механическое выравнивание уровней ряда
- Аналитическое выравнивание уровней ряда

Типы трендов

Линейный тренд $\hat{y}_t = a + b \cdot t$

Гипербола $\hat{y}_t = a + b / t$

Экспонента $\hat{y}_t = e^{a + b \cdot t^l}$

Степенной тренд $\hat{y}_t = a \cdot t^b$

Парабола k-го порядка $\hat{y}_t = a + b_1 \cdot t + b_2 \cdot t^2 + \ldots + b_k \cdot t^k$

Приемы выявления типа тенденции

- графически
- по абсолютным приростам и темпам роста сглаженных уровней
- метод последовательных разностей
- сравнительная оценка остаточной суммы квадратов и характеристик качества регрессии

Анализ структуры временного ряда

2 вопрос

Расчет значений сезонной компоненты методом скользящей средней

↓

Построение аддитивной или мультипликативной модели временного ряда

Как выбрать тип модели?

Амплитуда сезонных колебаний примерно постоянна

- Аддитивная модель временного ряда

Амплитуда колебаний возрастает или уменьшается

- Мультипликативная модель временного ряда

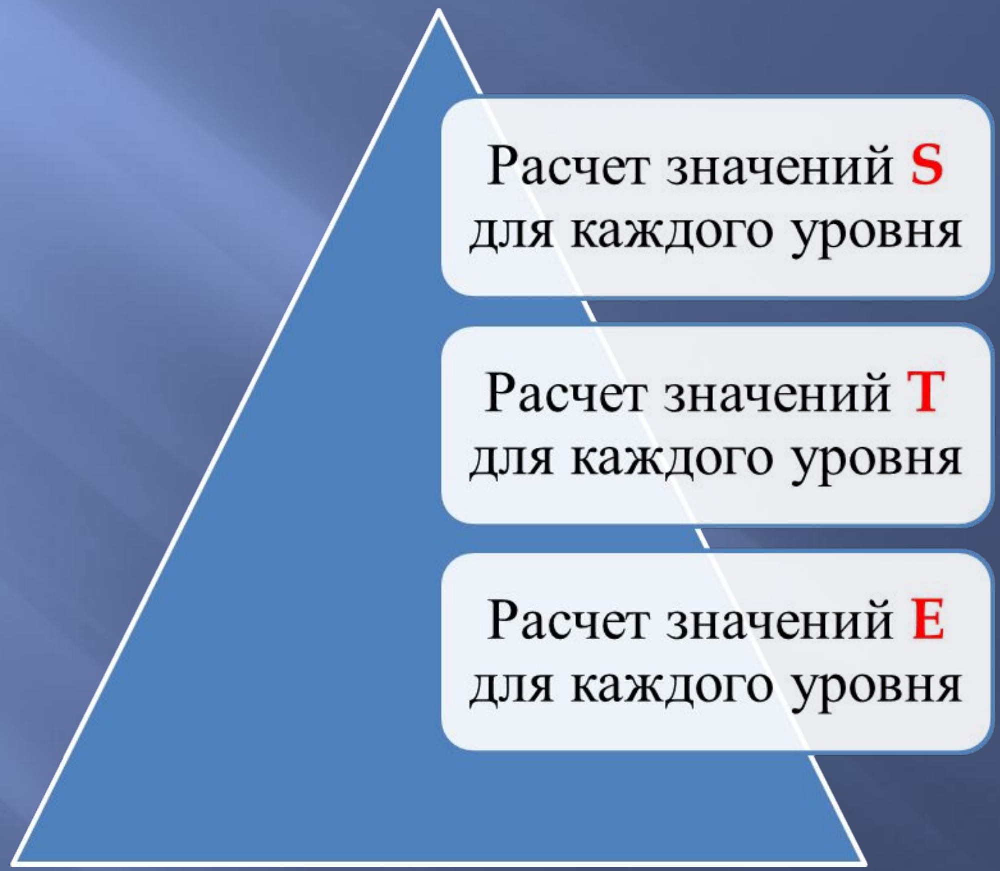
Аддитивная модель

- Сумма значений сезонной компоненты по всем кварталам равна нулю

Мультипликативная модель

- Сумма значений сезонной компоненты по всем кварталам равна числу периодов в цикле, то есть четырем

Процесс построения модели



Этапы построения модели

1. Выравнивание ряда
2. Расчет значений S
3. Расчет (T+E) или (T*E) через устранение S
4. Аналитическое выравнивание (T+E) или (T*E)
5. Расчет (T+S) или (T*S)
6. Расчет абсолютных E=Y-(T+S) и относительных ошибок

Построение аддитивной модели

1 шаг. Выравнивание уровней ряда

- Просуммируем уровни ряда за каждые четыре квартала со сдвигом на один момент времени
- Разделив полученные суммы на 4, найдем скользящие средние
- Найдем центрированные скользящие средние как средние значения из двух последовательных скользящих средних

Построение аддитивной модели

2 шаг.
Расчет сезонной компоненты S

- Найдем разность между уровнями и центрированными скользящими средними
- Расчет средней оценки сезонной компоненты для каждого квартала за все годы
- Расчет скорректированной сезонной компоненты

Моделирование сезонных колебаний

Аддитивная модель $Y_t = T_t + S_t + e_t$

Оценка сезонной компоненты за каждый квартал $s_t = y_t - \bar{y}_t$

Средняя оценка сезонной компоненты для квартала за все годы $\bar{S}_t = \frac{\sum s_t}{n}$

Скорректированная сезонная компонента $S_t = \bar{S}_t - k$

$$k = \frac{\sum_{t=1}^{4} \bar{S}_t}{4}$$

Построение аддитивной модели

3 шаг. Устранение сезонной компоненты S

- Вычтем скорректированное значение сезонной компоненты из каждого уровня исходного временного ряда
- Получим: T+E=Y-S

Построение аддитивной модели

4 шаг. Расчет значений тренда Т

- Проведем аналитическое выравнивание ряда (Т+Е) с помощью линейного тренда
- Рассчитаем значения Т для каждого момента времени по уравнению тренда

Построение аддитивной модели

5 шаг. Расчет значений T+S

- Прибавим к уровням T значения сезонной компоненты (S) для соответствующих кварталов

Построение аддитивной модели

6 шаг.
Расчет абсолютной ошибки

- Выполним расчет ошибки для каждого уровня ряда по формуле: E=Y-(T+S)
- Расчет суммы квадратов абсолютных ошибок и ее сравнение с общей суммой квадратов отклонений уровней ряда

Задача. Имеются поквартальные данные о потреблении электроэнергии в регионе за 4 года, в млн. квт.-час. Требуется построить аддитивную модель и найти прогнозную оценку потребления электроэнергии в 1 квартале следующего года.

Расчет сезонной компоненты S

№ квартала	Потребление электроэнер-гии	Итого за 4 квартала	Скользящая средняя за 4 квартала	Центрирован-ная скольз. средняя	Оценка сезонной компоненты
1	6,0	-	-	-	-
2	4,4	-	-	-	-
3	5,0	24,4	6,1	6,25	-1,250
4	9,0	25,6	6,4	6,45	2,550
5	7,2	26,0	6,5	6,625	0,575
6	4,8	27,0	6,75	6,875	-2,075
7	6,0	28,0	7,00	7,1	-1,100
8	10,0	28,8	7,20	7,3	2,700
9	8,0	29,6	7,40	7,45	0,550
10	5,6	30,0	7,50	7,625	-2,025
11	6,4	31,0	7,75	7,875	-1,475
12	11,0	32,0	8,00	8,125	2,875
13	9,0	33,0	8,25	8,325	0,675
14	6,6	33,6	8,4	8,375	-1,775
15	7,0	33,4	8,35	-	-
16	10,8	-	-	-	-

Расчет скорректированной сезонной компоненты S

Показатели	Год	№ квартала, i			
		1	2	3	4
	1	-	-	-1,250	2,550
	2	0,575	-2,075	-1,100	2,700
	3	0,550	-2,025	-1,475	2,875
	4	0,675	-1,775	-	-
Итого за i-й квартал (за все годы)		1,800	-5,875	-3,825	8,125
Средняя оценка сезонной компоненты для i-го квартала		0,600	-1,958	-1,275	2,708
Скорректированная сезонная компонента, Si		0,581	-1,977	-1,294	2,690

0,6-1,958-1,275+2,708=0,075.

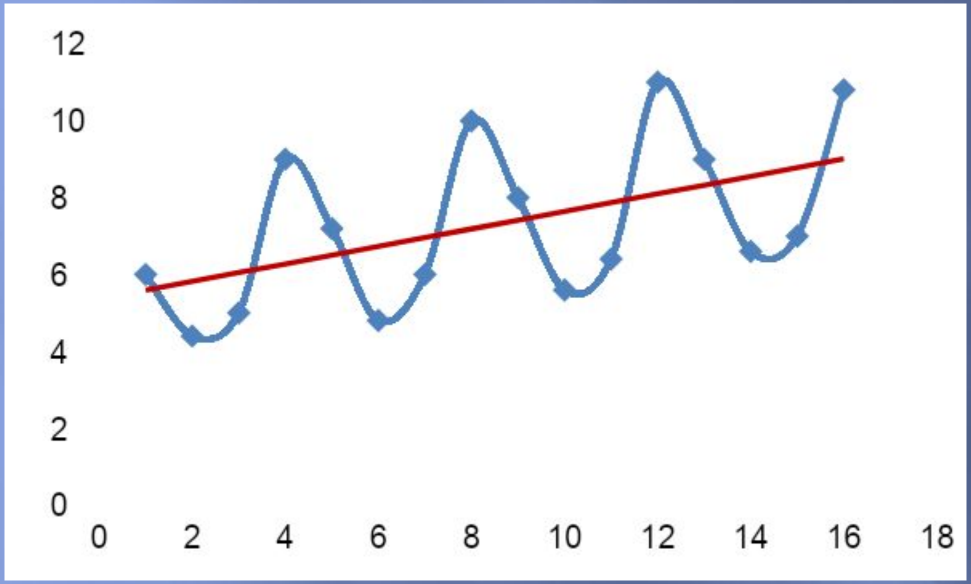
K=0,075/4=0,01875.

0,581-1,977-1,294+2,69=0.

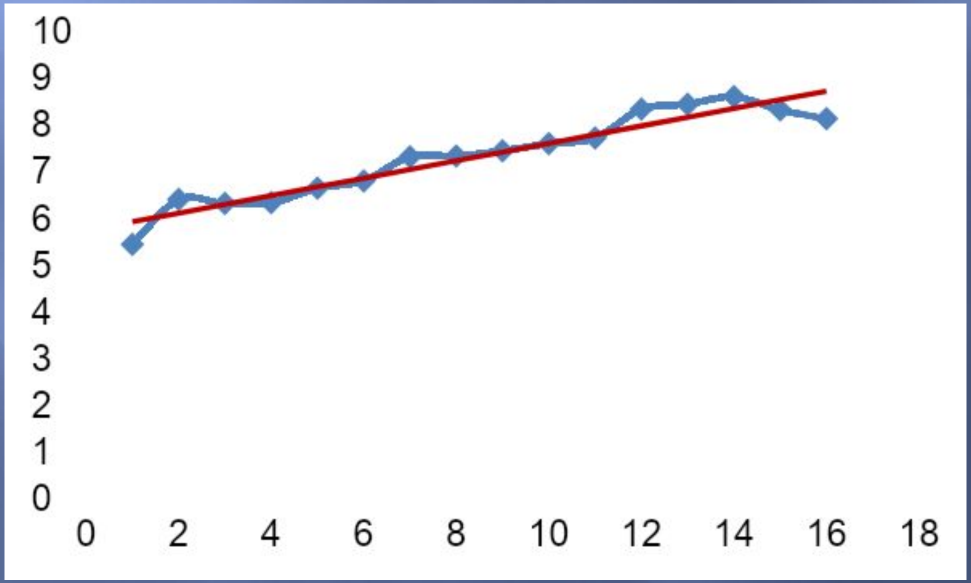
Расчет значений T+E и T+S

t	Yt	St	T+E= Yt-St	T	T+S	E=Yt-(T+S)	E^2
1	6,0	0,581	5,419	5,902	6,483	-0,483	0,2332
2	4,4	-1,977	6,377	6,088	4,111	0,289	0,0833
3	5,0	-1,294	6,294	6,275	4,981	0,019	0,0004
4	9,0	2,69	6,310	6,461	9,151	-0,151	0,0228
5	7,2	0,581	6,619	6,648	7,229	-0,029	0,0008
6	4,8	-1,977	6,777	6,834	4,857	-0,057	0,0032
7	6,0	-1,294	7,294	7,020	5,726	0,274	0,0749
8	10,0	2,69	7,310	7,207	9,897	0,103	0,0107
9	8,0	0,581	7,419	7,393	7,974	0,026	0,0007
10	5,6	-1,977	7,577	7,580	5,603	-0,003	0,0000
11	6,4	-1,294	7,694	7,766	6,472	-0,072	0,0052
12	11,0	2,69	8,310	7,952	10,642	0,358	0,1278
13	9,0	0,581	8,419	8,139	8,720	0,280	0,0785
14	6,6	-1,977	8,577	8,325	6,348	0,252	0,0634
15	7,0	-1,294	8,294	8,512	7,218	-0,218	0,0474
16	10,8	2,69	8,110	8,698	11,388	-0,588	0,3458

T=5,715+0,186*t, R^2=0,91.



Прогнозная оценка по
исходным данным:
Yt=0,2276*17+5,365=9,234



Прогнозная оценка по
аддитивной модели:
Tt=0,1864*17+5,7155=8,884
Yt=Tt+St=8,884+0,581=9,465

Построение мультипликативной модели

3 вопрос

1 шаг. Выравнивание уровней ряда

- Просуммируем уровни ряда за каждые четыре квартала со сдвигом на один момент времени
- Разделив полученные суммы на 4, найдем скользящие средние
- Найдем центрированные скользящие средние как средние значения из двух последовательных скользящих средних

Построение мультипликативной модели

2 шаг. Расчет сезонной компоненты S

- Найдем оценки сезонной компоненты как частное от деления уровней на центрированные скользящие средние
- Расчет средней оценки сезонной компоненты для каждого квартала за все годы
- Расчет скорректированной сезонной компоненты

Моделирование сезонных колебаний

Мультипликативная модель $Y_t = T_t \cdot S_t \cdot e_t$

Оценка сезонной компоненты за каждый квартал $s_t = \dfrac{y_t}{\bar{y}_t}$

Средняя оценка сезонной компоненты для квартала за все годы $\bar{S}_t = \dfrac{\sum s_t}{n}$

Скорректированная сезонная компонента $S_t = \bar{S}_t \cdot k$

$$k = \frac{4}{\sum_{t=1}^{4} \bar{S}_t}$$

Построение мультипликативной модели

3 шаг. Устранение сезонной компоненты S

- Разделим каждый уровень исходного временного ряда на скорректированное значение сезонной компоненты
- Получим: T*E=Y/S

Построение мультипликативной модели

4 шаг. Расчет значений тренда T

- Проведем аналитическое выравнивание ряда (T*E) с помощью линейного тренда
- Рассчитаем значения T для каждого момента времени по уравнению тренда

Построение мультипликативной модели

5 шаг. Расчет значений T+S

- Умножим уровни T на значения сезонной компоненты (S) для соответствующих кварталов

Построение мультипликативной модели

6 шаг. Расчет абсолютной ошибки

- Выполним расчет ошибки для каждого уровня ряда по формуле: E=Y/(T*S)
- Расчет суммы квадратов абсолютных ошибок и ее сравнение с общей суммой квадратов отклонений уровней ряда

Расчет сезонной компоненты S

№ квартала	Прибыль компании	Итого за 4 квартала	Скользящая средняя за 4 квартала	Центрирован-ная скол. средняя	Оценка сезонной компоненты
1	72	-	-	-	-
2	100	-	-	-	-
3	90	326	81,5	81,250	1,108
4	64	324	81,0	80,000	0,800
5	70	316	79,0	77,750	0,900
6	92	306	76,5	75,750	1,215
7	80	300	75,0	74,000	1,081
8	58	292	73,0	71,500	0,811
9	62	280	70,0	68,500	0,905
10	80	268	67,0	65,750	1,217
11	68	258	64,5	63,250	1,075
12	48	248	62,0	59,500	0,807
13	52	228	57,0	54,750	0,950
14	60	210	52,5	50,250	1,194
15	50	192	48,0	-	-
16	30	-	-	-	-

Расчет скорректированной сезонной компоненты S

Показатели	Год	№ квартала, i			
		1	2	3	4
	1	-	-	1,108	0,800
	2	0,900	1,215	1,081	0,817
	3	0,905	1,217	1,075	0,807
	4	0,950	1,194	-	-
Итого за i-й квартал (за все годы)		2,755	3,626	3,264	2,424
Средняя оценка сезонной компоненты для i-го квартала		0,918	1,209	1,088	0,808
Скорректированная сезонная компонента, Si		0,913	1,202	1,082	0,803

0,918+1,209+1,088+0,808=4,023

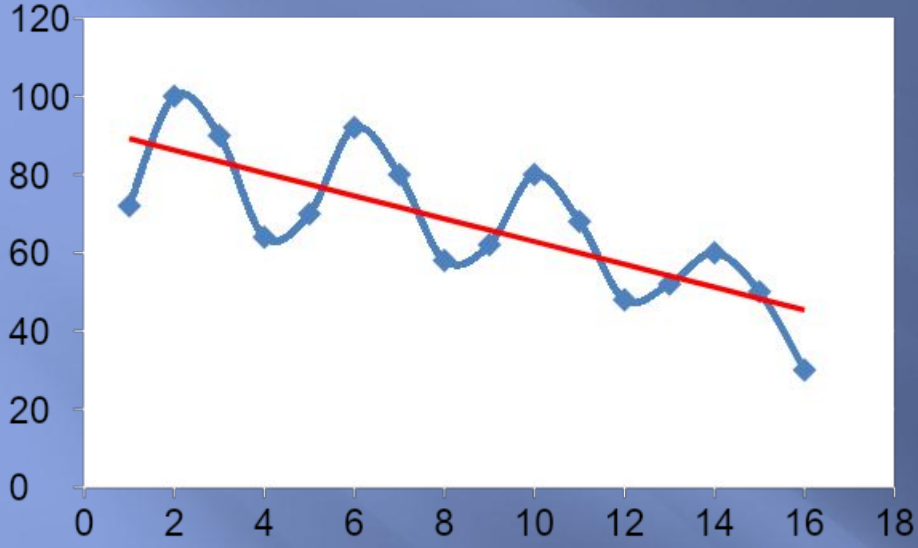
K=4/4,023=0,9943.

0,913+1,202+1,082+0,803=4.

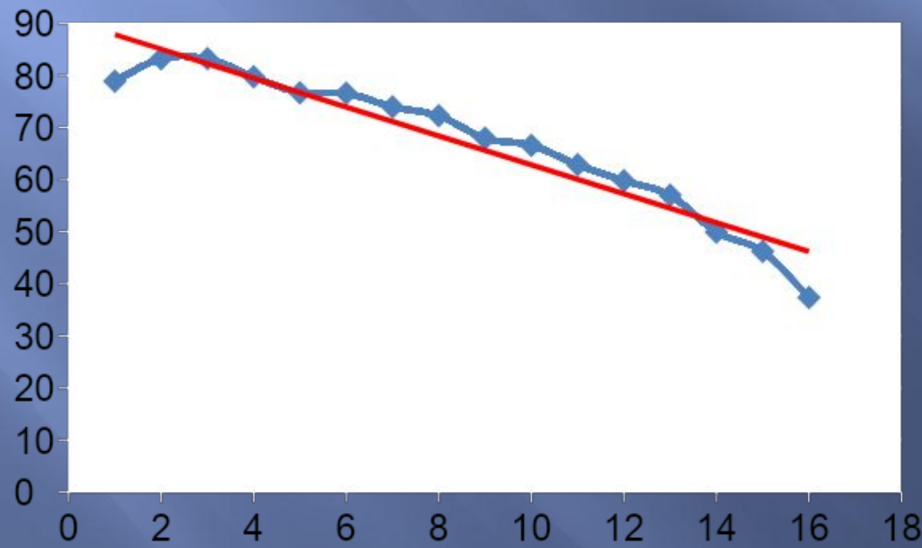
Расчет значений T*E и T*S

t	Yt	St	T*E=Yt/St	T	T*S	E=Yt/(T*S)	E^2
1	72	0,913	78,86	87,80	80,16	0,898	66,66
2	100	1,202	83,19	85,03	102,2	0,978	4,86
3	90	1,082	83,18	82,25	89,00	1,011	1,00
4	64	0,803	79,70	79,48	63,82	1,003	0,03
5	70	0,913	76,67	76,70	70,03	1,000	0,00
6	92	1,202	76,54	73,93	88,86	1,035	9,85
7	80	1,082	73,94	71,15	76,99	1,039	9,08
8	58	0,803	72,23	68,38	54,91	1,056	9,57
9	62	0,913	67,91	65,60	59,90	1,035	4,43
10	80	1,202	66,56	62,83	75,52	1,059	20,08
11	68	1,082	62,85	60,05	64,98	1,047	9,14
12	48	0,803	59,78	57,28	45,99	1,044	4,03
13	52	0,913	56,96	54,50	49,76	1,045	5,02
14	60	1,202	49,92	51,73	62,18	0,965	4,73
15	50	1,082	46,21	48,95	52,97	0,944	8,79
16	30	0,803	37,36	46,18	37,08	0,809	50,12

$T=90,59-2,773*t$, $R^2=0,92$.



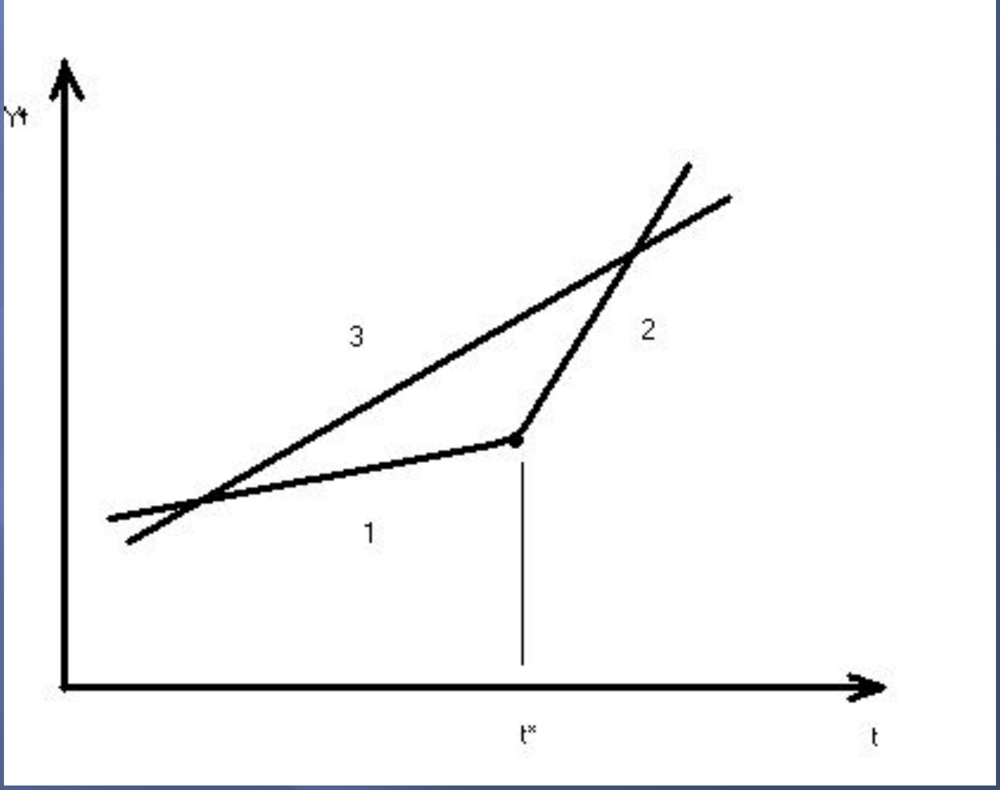
Прогнозная оценка по
исходным данным:
Yt=-2,9235*17+92,1=42,401



Прогнозная оценка по
мультипликативной модели:
Tt=-2,7749*17+90,578=43,405
Yt=Tt*St=43,405*0,913=39,628

Моделирование тенденции временного ряда при наличии структурных изменений

Кусочно-линейная модель регрессии (тренда)



Тест Чоу

Остаточная сумма квадратов по кусочно-линейной модели

$$S^2 = S^2{}_1 + S^2{}_2$$

Число степеней свободы по кусочно-линейной модели

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

Уменьшение остаточной дисперсии при переходе от единого тренда к кусочно-линейной модели

$$\Delta S^2 = S^2{}_t - S^2$$

Число степеней свободы для уменьшения остаточной дисперсии

$$n - k_{t3} - (n - k_1 - k_2) = k_1 + k_2 - k_{t3}$$

Тест Чоу

Фактическое значение F-критерия:

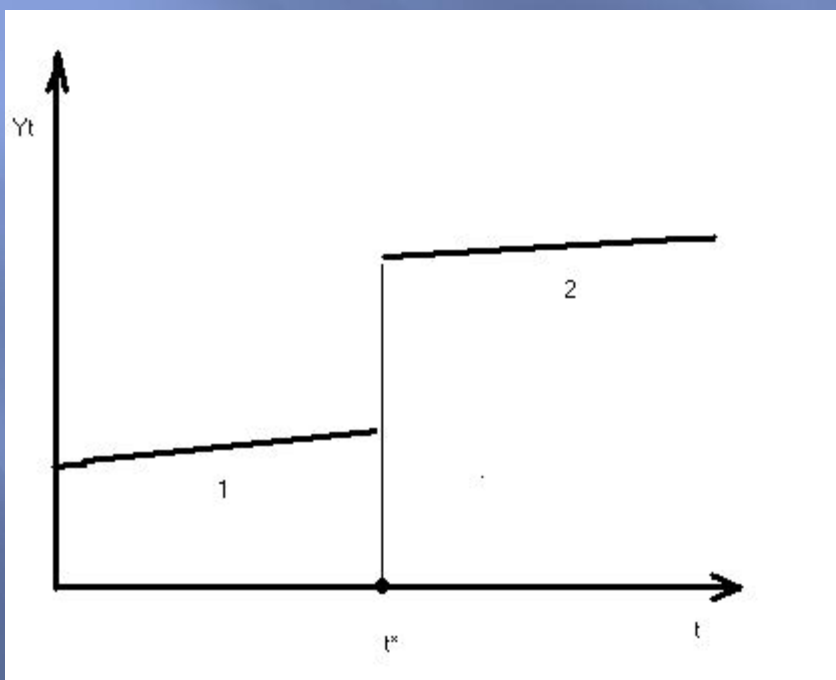
$$F = \frac{\Delta S^2 / (k_1 + k_2 - k_{t3})}{S^2 / (n - k_1 - k_2)}$$

$$F\alpha ; v_1 = k_1 + k_2 - k_{t3} ; v_2 = n - k_1 - k_2$$

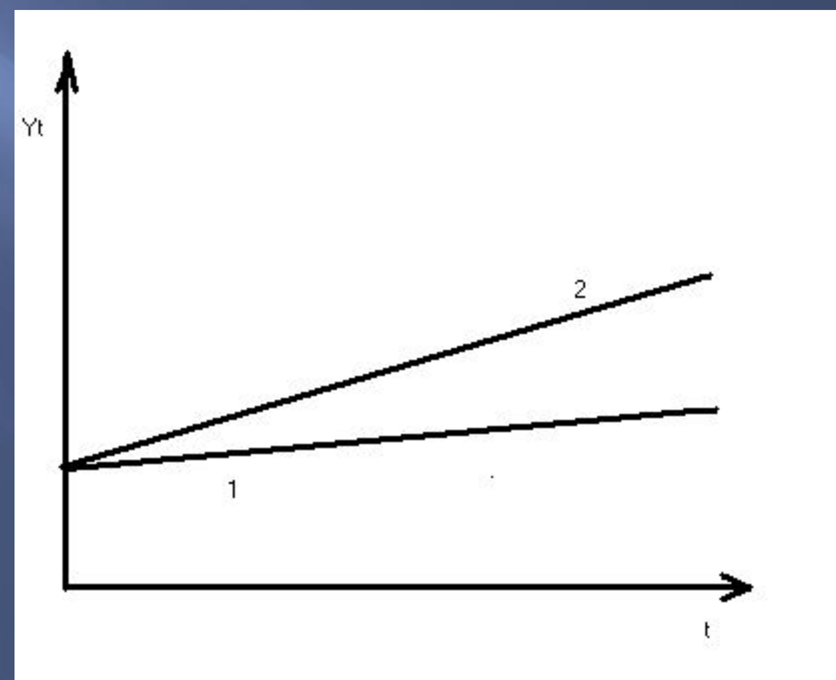
Применяется кусочно-линейная модель, если Fнабл>Fтабл

Частные случаи кусочно-линейной модели

$a_2=a_1$, различие между b_2 и b_1 статистически незначимо



$b_2=b_1$, различие между a_2 и a_1 статистически незначимо



Проблемы при изучении взаимосвязи временных рядов

- устранение сезонной и циклической компоненты
- завышенный парный коэффициент корреляции
- автокорреляция остатков
- мультиколлинеарность факторов

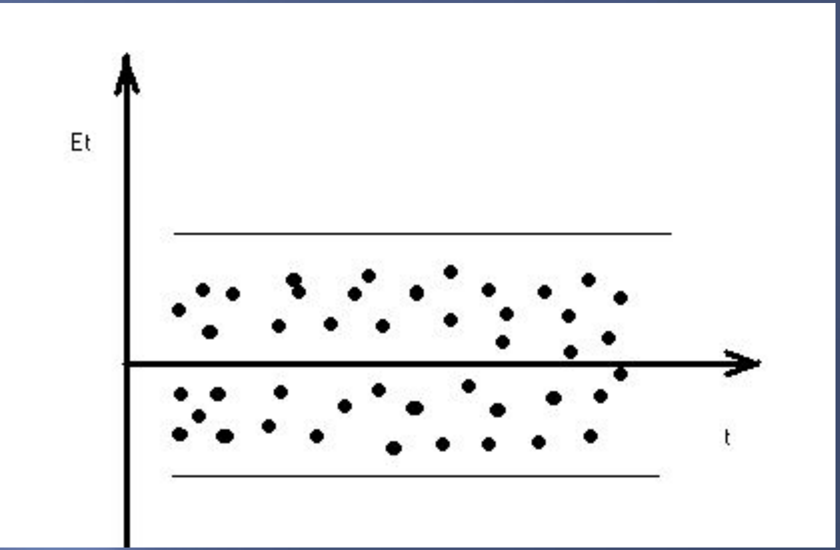
Методы исключения тенденции

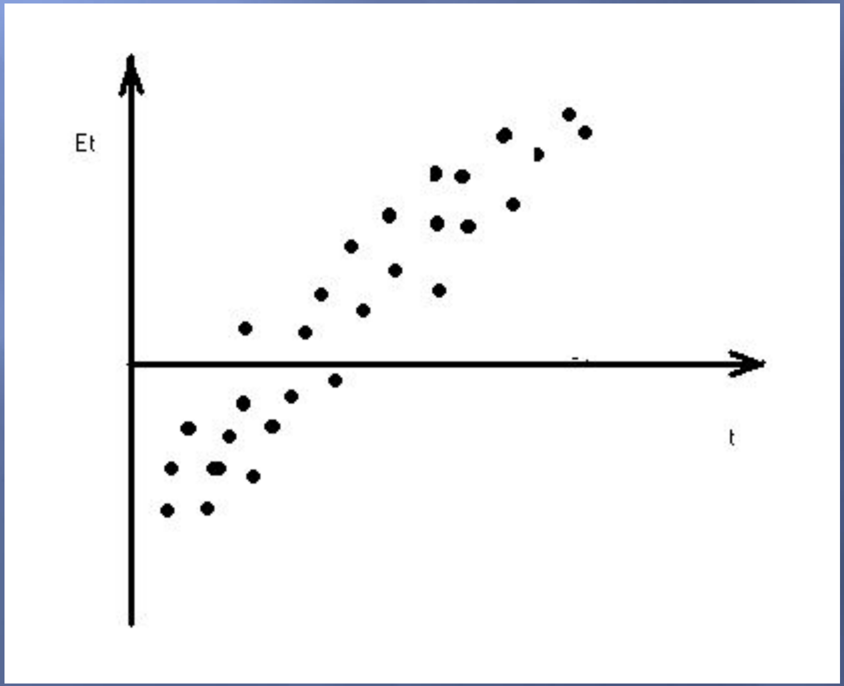
- Преобразование уровней ряда в новые переменные (метод последовательных разностей, метод отклонений от трендов)
- Элиминирование влияния фактора времени на Yt и Xt (метод включения в модель регрессии фактора времени)

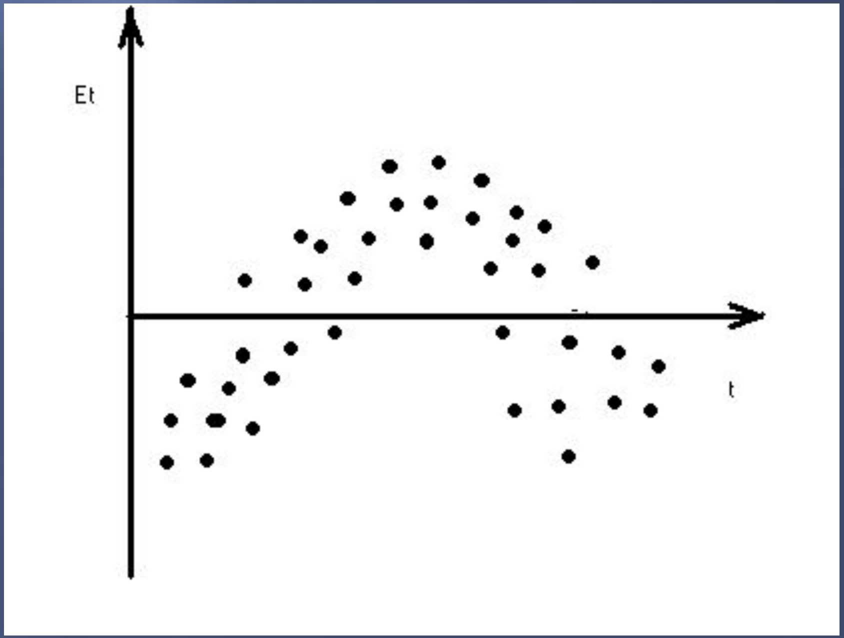
Автокорреляция в остатках уровней временного ряда

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

$$\hat{y}_t = a + \sum_{j=1}^{k} b_j \cdot x_{jt} + \varepsilon_t$$







Методы выявления автокорреляции остатков

Критерий Дарбина-Уотсона

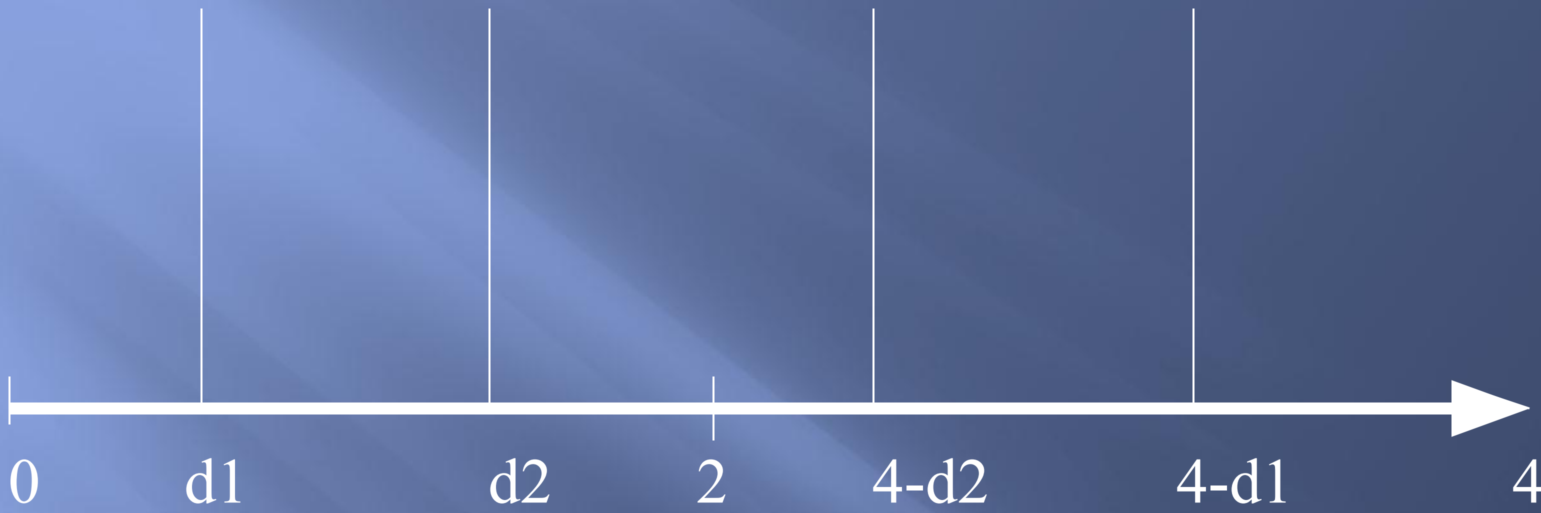
$$DW = \frac{\sum_{t=2}^{n} (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{n} \varepsilon_t^{\ 2}}$$

Коэффициент автокорреляции остатков 1 порядка

$$r_1^{\ \varepsilon} = \frac{\sum_{t=2}^{n} \varepsilon_t \cdot \varepsilon_{t-1}}{\sum_{t=2}^{n} \varepsilon_t^{\ 2}}$$

$$DW \approx 2 \cdot (1 - r_1^{\ \varepsilon})$$

Проверка гипотезы о наличии автокорреляции остатков



Положи-
тельная
автокор.

Зона
неопре-
делен.

Отсутствие
автокорре-
ляции

Зона
неопре-
делен.

Отрица-
тельная
автокор.

Доверительный интервал точечного прогноза

$$\hat{y}_t = a + b \cdot x_t$$

$$\hat{y}_{t+l} - t_{\alpha/2,n-1} \cdot S_{(\Delta t+l)} <$$

$$< y_{t+l} < \hat{y}_{t+l} + t_{\alpha/2,n-1} \cdot S_{(\Delta t+l)}$$

$$S_{(\Delta t+l)} = \sqrt{(1 + \frac{1}{n} + \frac{(x_{t+l} - \bar{x})^2}{n \cdot S^2{}_x}) \cdot S_e{}^2}$$

$$S^2{}_x = \sum (x_t - \bar{x})^2$$

$$S_e{}^2 = \frac{\sum e_t{}^2}{n-2}$$