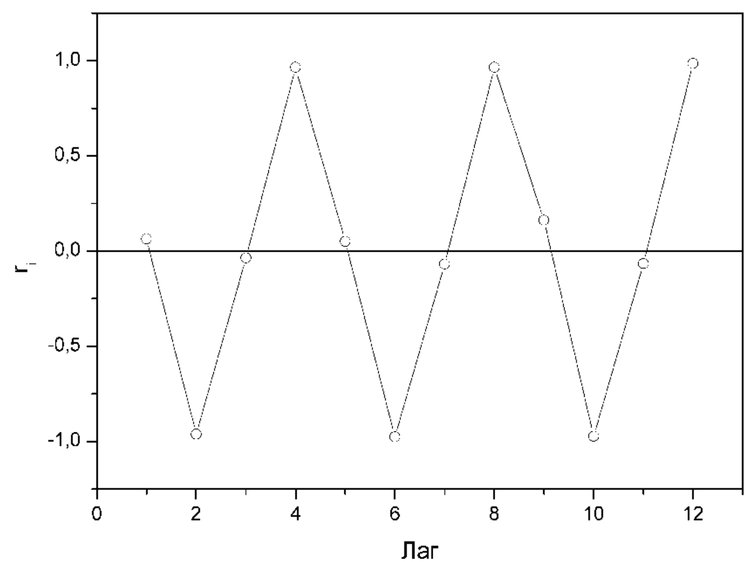
АВТОКОРРЕЛЯЦИЯ УРОВНЕЙ ВРЕМЕННОГО РЯДА



Временные ряды

Эконометрическую модель можно построить, используя два типа исходных данных:

- данные, характеризующие совокупность различных объектов в определенный момент (период) времени;
- данные, характеризующие один объект за ряд последовательных моментов (периодов) времени.

Модели, построенные по данным первого типа, называются ***пространственными моделями***. Модели, построенные по данным второго типа, называются ***моделями временных рядов.***

Временной ряд (динамический ряд, ряд динамики) – это совокупность значений какого-либо показателя за несколько последовательных моментов (периодов) времени.

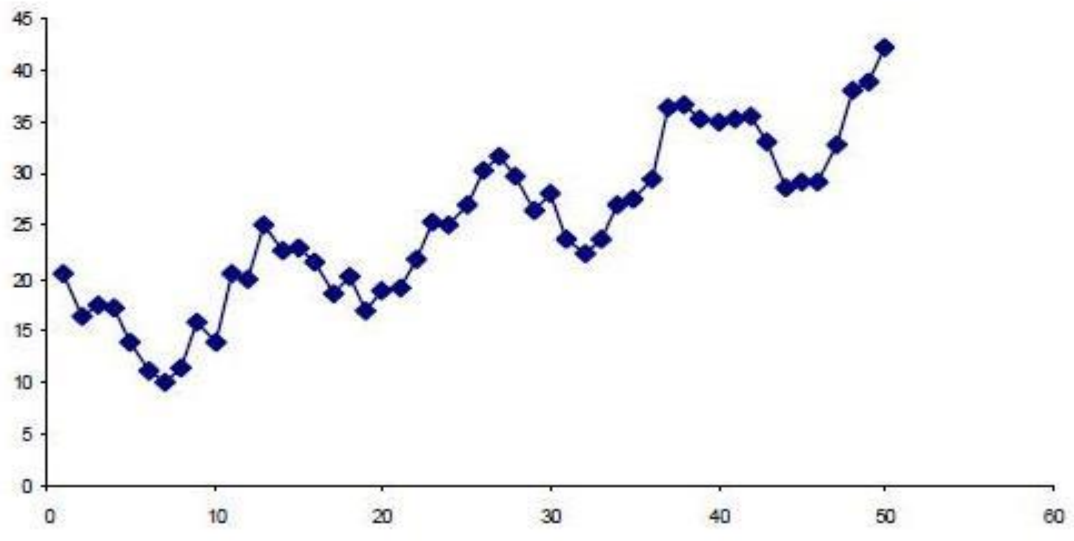
	2000 г.	2001 г.	2002 г.	2003 г.	2004 г.
ВВП, млрд. руб.	7305,6	8943,6	10834,2	13285,2	17048,1

Три составляющие временного ряда

Долговременная тенденция **T**

Случайная компонента **E**

Периодические (циклические или сезонные) колебания **S**



Модели временного ряда:

1) аддитивная	$Y_t = T_t + S_t + E_t$
2) мультипликативная	$Y_t = T_t \times S_t \times E_t$
3) смешанная	$Y_t = T_t \times S_t + E_t$

Основная задача эконометрического исследования временного ряда:

выявление и количественное выражение его компонент (тенденции, периодичности, случайной компоненты)

в целях их использования для прогнозирования будущих значений ряда.

Автокорреляция уровней временного ряда –

это корреляционная зависимость между последовательными уровнями временного ряда.

Измеряется с помощью *линейного коэффициента корреляции* между уровнями исходного временного ряда и уровнями ряда, сдвинутыми на несколько шагов назад во времени:

$$r_\tau = \frac{\sum_{t=\tau+1}^{n}(y_t - \bar{y}_{1\tau})\cdot(y_{t-\tau} - \bar{y}_{2\tau})}{\sqrt{\sum_{t=\tau+1}^{n}(y_t - \bar{y}_{1\tau})^2 \cdot \sum_{t=\tau+1}^{n}(y_{t-\tau} - \bar{y}_{2\tau})^2}}$$

$$\bar{y}_{1\tau} = \frac{\sum_{t=\tau+1}^{n} y_t}{n-\tau} \qquad \bar{y}_{2\tau} = \frac{\sum_{t=\tau+1}^{n} y_{t-\tau}}{n-\tau}$$

τ – величина сдвига во времени, или лаг

Например, лаг τ=1 означает, что ряд сдвинут на один период (момент) назад и т.д. С увеличением лага число пар значений, по которым рассчитывается коэффициент автокорреляции, уменьшается.

τ=1 =>

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

τ=2 =>

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

Свойства коэффициента автокорреляции:

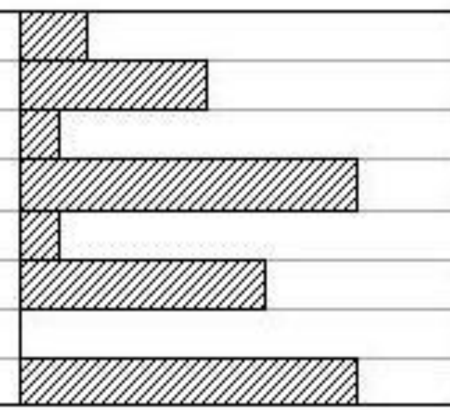
- характеризует *тесноту только линейной связи* текущего и предыдущего уровней ряда, поэтому по данному коэффициенту можно судить о наличии линейной или близкой к линейной тенденции. Для некоторых временных рядов, имеющих сильную нелинейную тенденцию, коэффициент автокорреляции может приближаться к нулю;

- по *знаку* коэффициента автокорреляции нельзя судить о возрастающей или убывающей тенденции в уровнях ряда.

Автокорреляционная функция временного ряда (АКФ) *– это последовательность коэффициентов автокорреляции первого, второго и т.д. порядков.*
Коррелограмма – *это график зависимости значений АКФ от величины лага.*

Коррелограмма временного ряда потребления электроэнергии

Лаг (квартал)	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	



Моделирование тенденции временного ряда

Аналитическое выравнивание – это построение аналитической функции, характеризующей зависимость уровней ряда от времени, т.е. построение тренда:

- *линейный тренд* $\hat{y}_t = a + bt$
- *экспоненциальный тренд* $\hat{y}_t = e^{a+bt}$
- *гипербола* $\hat{y}_t = a + b/t$
- *тренд в форме степенной функции*

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

Для определения вида тенденции применяются следующие методы:

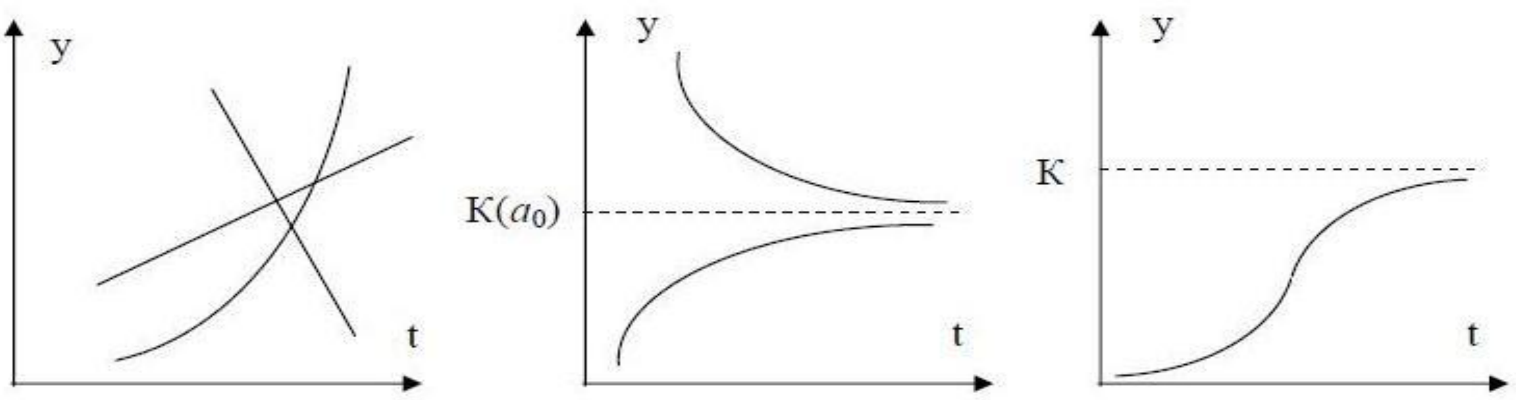
– качественный анализ изучаемого процесса;

– построение и визуальный анализ графика зависимости уровней ряда от времени;

– расчет и анализ показателей динамики временного ряда (абсолютные приросты, темпы роста и др.);

– метод перебора, при котором строятся тренды различного вида с последующим выбором наилучшего на основании значения скорректированного коэффициента детерминации.

Выбор вида тенденции на основе качественного анализа



а) I класс *б)* II класс *в)* III класс

Процессы с монотонным характером развития и отсутствием пределов роста

Функции:

✓линейная,
✓параболическая,
✓экспоненциальная,
✓степенная.

Процессы, имеющие предел роста (падения), так называемые процессы с «насыщением»

Функции:

✓гиперболическая,
✓модифицированная экспонента.

S-образные процессы

Функция:

✓логистическая.

$$y_t = \frac{K}{1 + a_0 e^{-bt}}$$

Моделирование периодических колебаний

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

Процесс построения модели включает в себя следующие этапы:

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

1 этап. Выравнивание исходного ряда методом скользящей средней

Расчет оценок сезонной компоненты в аддитивной модели

Кварталы	Потребление эл/энергии	Итого за 4 квартала	Скользящая средняя за 4 квартала	Центрированная скользящая средняя	Оценка сезонной компоненты
1	*2*	*3*	*4*	*5*	*6=2-5*
1	6,0				
2	4,4	24,4	6,10		
3	5,0	25,6	6,40	6,250	-1,250
4	9,0	26,0	6,50	6,450	2,550
5	7,2	27,0	6,75	6,625	0,575
6	4,8	28,0	7,00	6,875	-2,075
7	6,0	28,8	7,20	7,100	-1,100
8	10,0	29,6	7,40	7,300	2,700
9	8,0	30,0	7,50	7,450	0,550
10	5,6	21,0	7,75	7,625	-2,025
11	6,4	32,0	8,00	7,875	-1,475
12	11,0	33,0	8,25	8,125	2,875
13	9,0	33,6	8,40	8,325	0,675
14	6,6	33,4	8,35	8,375	-1,775
15	7,0				
16	10,8				

2 этап. Расчет значений периодической компоненты S

Расчет значений сезонной компоненты в аддитивной модели

Показатель	Год	Кварталы			
		1	2	3	4
	1ый	-	-	-1,250	2,550
	2ой	0,575	-2,075	1,100	2,700
	3ий	0,550	-2,025	-1,475	2,875
	4ый	0,675	-1,775	-	-
Итого за i-й квартал (за все годы)	x	1,800	-5,875	-3,825	8,125
Средняя оценка сезонной компоненты для i-го квартала, $\bar{S}_i$	x	0,600	-1,958	-1,275	2,708
Скорректированная сезонная компонента, S_i	x	0,581	-1,977	-1,294	2,690

3 этап. Устранение периодической компоненты из исходных уровней ряда и получение выравненных данных *(T+E)*

Расчет выравненных значений T и E в аддитивной модели

t	y	S	$T+E=$ $y-S$	T	$T+S$	$E=$ $y-(T+S)$	E^2
1	2	3	4	5	6	7	8
1	6,0	0,581	5,914	5,902	6,483	-0,483	0,2333
2	4,4	-1,977	6,337	6,088	4,111	0,289	0,0835
3	5,0	-1,294	6,294	6,275	4,981	0,019	0,0004
4	9,0	2,690	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	5,727	0,273	0,0745
8	10,0	2,690	7,310	7,207	9,896	0,104	0,0108
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	-0,030	0,0009
11	6,4	-1,294	7,694	7,766	6,472	-0,072	0,0052
12	11,0	2,690	8,310	7,952	10,642	0,358	0,1282
13	9,0	0,581	8,419	8,139	8,720	0,280	0,0784
14	6,6	-1,977	8,577	8,325	6,348	0,252	0,0635
15	7,0	-1,294	8,294	8,519	7,218	-0,218	0,0475
16	10,8	2,690	8,110	8,698	11,388	-0,588	0,3457

4 этап. Аналитическое выравнивание уровней ряда и расчет значений *T* с использованием полученного уравнения тренда

$$T = 5{,}715 + 0{,}186t$$

Моделирование сезонных и циклических колебаний

- Два подхода
  - Расчет сезонной компоненты методом скользящей средней и построение аддитивной или мультипликативной модели;
  - Применение фиктивных переменных.
- Аддитивная модель
$$Y=T+S+E$$
- Мультипликативная модель
$$Y=TSE$$
- T - трендовая составляющая,
- S – циклическая (сезонная) составляющая,
- E – случайная составляющая.

Алгоритм построения модели(методом скользящей средней)

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

	аддитивная модель	мультипликативная модель
корректирующий коэффициент:	$f = \frac{\sum \overline{S}_i}{n}$	$f = \frac{n}{\sum \overline{S}_i}$
корректированные сезонные компоненты	$S_i = \overline{S_i} - f,$	$S_i = \overline{S}_i \cdot f$
контроль	сумма значений сезонной компоненты равна 0	сумма значений сезонной компоненты = числу периодов в цикле
выравнивание сезонной компоненты	$T + E = Y - S$	$T \cdot E = Y / S$
расчет ошибок	$E = Y - (T + S)$	$E = y / (T \cdot S)$
абсолютные ошибки		E' = y_t-(T*S)
оценка качества модели	$1 - \frac{\sum E^2}{\sum (y_t - \bar{y}_t)^2}$	$1 - \frac{\sum E'^2}{\sum (y_t - \bar{y}_t)^2}$

- **Применение фиктивных переменных для моделирования сезонных колебаний**

Количество фиктивных переменных в такой модели должно быть на единицу меньше числа моментов (периодов) времени внутри одного цикла колебания.

Пусть имеется временной ряд, содержащий циклические колебания периодичностью k.

Модель регрессии с фиктивными переменными для этого ряда будет иметь вид:

$$y_t = a + b \cdot t + c_1 x_1 + ... + c_j x_j + \ldots + c_{k-1} x_{k-1} + \varepsilon_t \qquad (*)$$

$$где\ x_j = \begin{cases} 1\ для\ каждого\ j\ внутри\ каждого\ цикла, \\ 0\ во\ всех\ остальных\ случаях. \end{cases}$$

- Параметр ***b*** в этой модели характеризуют среднее абсолютное изменение уровней ряда под воздействием тенденции.

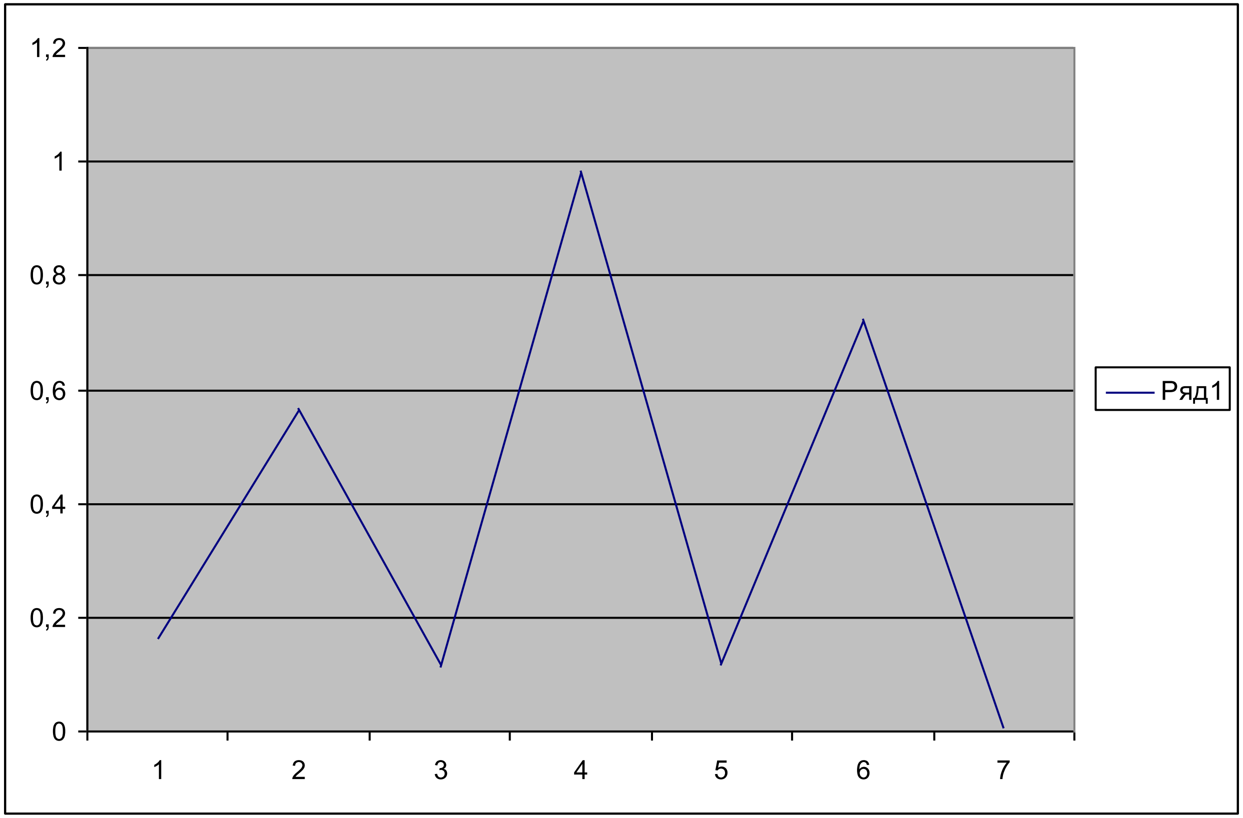
- В сущности, модель (*) - аналог аддитивной модели временного ряда.

Пример

- Построим модель регрессии с включением фактора времени и фиктивных переменных для данных о потреблении электроэнергии за 16 кварталов, млн.кВт.ч.

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

r1	0,165155
r2	-0,56687
r3	0,113558
r4	0,983025
r5	0,118711
r6	-0,72205
r7	-0,00337



- Составим матрицу исходных данных

t	x_1	x_2	x_3	y
1	1	0	0	6,0
2	0	1	0	4,4
3	0	0	1	5,0
4	0	0	0	9,0
5	1	0	0	7,2
6	0	1	0	4,8
7	0	0	1	6,0
8	0	0	0	10,0
9	1	0	0	8,0
10	0	1	0	5,6
11	0	0	1	6,4
12	0	0	0	11,0
13	1	0	0	9,0
14	0	1	0	6,6
15	0	0	1	7,0
16	0	0	0	10,8

- Оценим параметры уравнения регрессии (*) обычным МНК. Результаты оценки приведем в табл.

переменная	коэффициент	t-критерий
Const t x_1 x_2 x_3	8.3250 0.1875 -2.0875 -4.4750 -3.9125	36.6318 11.0691 -9.4797 -20.6292 -18.2034
R^2 =0,985		t табл = 2

- Уравнение регрессии имеет вид:

$$\hat{y}_t = 8.33 + 0.19t - 2.09 x_1 - 4.48 x_2 - 3.91 x_3.$$

- Влияние сезонной компоненты в каждом квартале статистически значимо ($t_{крит}$=2).
- Сезонные колебания в I, II, III кварталах приводят к снижению этой величины.
- В уровнях ряда присутствует возрастающая тенденция.

КОИНТЕГРАЦИЯ ВРЕМЕННЫХ РЯДОВ

Коинтеграцией называется зависимость в уровнях двух (или более) временных рядов, которая выражается в совпадении или противоположной направленности их тенденций.

- Рассмотрим уравнение регрессии вида:

$$y_t = a + bx_t + \varepsilon_t \qquad (**)$$

- Одним из методов тестирования гипотезы о коинтеграции временных рядов y_t и x_t является **критерий Энгеля-Грангера.**

- Алгоритм применения критерия:

- 1. Выдвигается ноль-гипотеза об отсутствии коинтеграции между рядами x_t и y_t.
- 2.Рассчитывают параметры уравнения регрессии вида (**)
- Где $\Delta\varepsilon_t$-первые разности остатков, полученных из соотношения
- 3.Определяют фактическое значение *t*-критерия для коэффициента a регрессии в уравнении (**).

$$\Delta\varepsilon_t = a + b \cdot \varepsilon_{t-1}.$$

$$\varepsilon_t = y_t - a - b \cdot x_t.$$

- 4.Сравнивают полученное значение с критическим значением статистики τ.
- Если фактическое значение t больше критического значения τ для заданного уровня значимости α, то ноль-гипотезу об отсутствии коинтеграции исследуемых временных рядов отклоняют и с вероятностью $(1-\alpha)$ принимают альтернативную гипотезу о том, что между рядами есть коинтеграция.

- В противном случае гипотеза об отсутствии коинтеграции между исследуемыми рядами не отклоняется.

Поскольку коинтеграция означает совпадение динамики временных рядов в течение длительного промежутка времени, то сама эта концепция применима только к временным рядам, охватывающим сравнительно длительные (например, в несколько десятилетий) промежутки времени.

- **Пример**. Пусть имеются данные о среднедушевом располагаемом доходе и среднедушевом расходе на конечное потребление в США в период с 1960 по 1991 г.
- Провести тестирование временных рядов среднедушевого дохода и расхода на потребление на коинтеграцию.

Год,	Среднедушевой располагаемый доход	Среднедушевые расходы на конечное потребление	Остатки	Скорректированные остатки	
				дохода,	расхода,
1	2	3	4	5	6
1960	7264	6698	173,80	-	-
1961	7382	6740	106,98	2092,87	1862,99
1962	7583	6931	112,61	2207,95	2023,41
1963	7718	7089	146,12	2196,60	2042,34
1964	8140	7384	51,94	2520,30	2222,29
1965	8508	7703	31,57	2581,03	2326,50
1966	8822	8005	43,99	2627,08	2396,22
1967	9114	8163	-67,29	2690,45	2334,33
1968	9399	8506	12,88	2762,83	2562,28
1969	9606	8737	52,98	2762,32	2543,54
1970	9875	8842	-90,10	2880,59	2480,34
1971	10111	9022	-127,74	2920,73	2583,88
1972	10414	9425	-4,17	3051,89	2855,82
1973	11013	9752	-229,57	3430,27	2889,38
1974	10832	9602	-212,65	2813,12	2501,29
1975	10906	9711	-171,90	3018,91	2719,51
1976	11192	10121	-25,65	3251,03	3050,14
1977	11406	10425	81,00	3256,78	3055,61
1978	11851	10744	-10,39	3545,96	3153,26
1979	12039	10867	-51,76	3409,94	3052,98

Год,	Среднедушевой располагаемый доход (долл.США),	Среднедушевые расходы на конечное потребление	Остатки,	Скорректированные остатки	
				дохода,	расход а,
1	2	3	4	5	6
1980	12005	10746	-150,41	3239,06	2826,87
1981	12156	10770	-265,66	3414,81	2945,43
1982	12146	10782	-244,44	3294,86	2940,05
1983	12349	11179	-34,65	3505,15	3328,31
1984	13029	11617	-223,75	4037,34	3477,25
1985	13258	12015	-36,94	3771,21	3556,33
1986	13552	12336	112,93	3898,47	3587,53
1987	13545	12568	251,39	3677,40	3585,80
1988	13890	12903	268,22	4027,49	3751,88
1989	14030	13027	263,11	3916,29	3631,95
1990	14154	13051	172,76	3938,35	3565,66
1991	13987	12889	164,77	3681,06	3386,19

- Регрессионный анализ зависимости среднедушевых расходов на конечное потребление от среднедушевого располагаемого дохода показал следующее:
- Константа -174,746
- Коэффициент регрессии 0,922212
- Стандартная ошибка 0,012837
- R-квадрат 0,994221
- Число наблюдений 32
- Число степеней свободы 30
- Уравнение регрессии имеет вид:

$$\hat{y}_t = -174{,}75 + 0{,}922 \cdot x_t + \varepsilon_t .$$

- Применим критерий Энгеля-Грангера. Воспользовавшись полученным уравнением регрессии,определим остатки (см. табл.). Определим параметры уравнения регрессии:
-
- Константа -1,7293
- Коэффициент регрессии -0,2724
- Стандартная ошибка 0,126806
- R-квадрат 0,137319
- Число наблюдений 31
- Число степеней свободы 30

$$\Delta\varepsilon_t = a + b \cdot \varepsilon_{t-1}.$$

- Фактическое значение *t*-критерия, рассчитанное по данным уравнения регрессии,равно **-2,154**.
- Критическое значение $\tau_{0,05}$ =**1,9439**

- Вывод: с вероятностью 95% можно отклонить ноль-гипотезу и сделать вывод о коинтеграции временных рядов среднедушевого дохода и среднедушевых расходов на конечное потребление