Тема 1. Обробка вхідного масиву економічних даних

План

1.1 Методи моделювання часових рядів.

1.2 Основні поняття і попередній аналіз рядів динаміки.

1.3Основні характеристики динаміки часового ряду.

1.4 Систематичні та випадкові компоненти часового ряду.

1.5 Перевірка на аномальність (стаціонарність, однорідність).

1.6 Нормалізація показників.

Поняття часового ряду

$$y = y_1, y_2 \boxtimes \; y_n$$

$$t_1, t_2 \boxtimes \; t_n$$

$$y_t \quad t = 1, 2, \ldots, n$$

Фактори формування рівнів часового ряду

Фактори формування тенденції ряду

Фактори формування циклічних коливань ряду

Випадкові фактори

Моментний часовий ряд

Дата надання позички	01.10.	05.10.	12.10.	23.10.	03.11.	07.11.
Розмір наданої позички, тис. гр. од.	3747	3710	3839	3783	3747	3710

Інтервальний часовий ряд

Місяць	Січень	Лютий	Березень	Квітень	Травень
Валовий внутрішній продукт, млн. грн.	6578	7016	7353	7353	7941

Часовий ряд, утворений із середніх значень показника

Місяць	Січень	Лютий	Березень	Квітень	Травень
Середня зарплата загалом, грн/міс.	152,2	153,7	165,8	161,6	163,71

Основні характеристики динаміки часового ряду

Характеристики	Розрахункові формули
1. Абсолютний приріст	$\Delta y_i = y_i - y_{i-1}$
2. Коефіцієнт зростання	$K_{i(зр)} = \dfrac{y_i}{y_{i-1}}$
3. Коефіцієнт приросту	$K_{i(пр)} = \dfrac{y_i - y_{i-1}}{y_{i-1}}$
4. Темп зростання	$T_{i(зр)} = \dfrac{y_i}{y_{i-1}} \cdot 100\% = K_{i(зр)} \cdot 100\%$
5. Темп приросту	$T_{i(пр)} = T_{i(зр)} - 100\ \%$, або $T_{i(пр)} = \dfrac{y_i - y_{i-k}}{y_{i-k}} \cdot 100\%$

6. Середня арифметична	$\bar{y} = \frac{\sum y_i}{n}$
7. Середня хронологічна	$\bar{y} = \frac{\sum y_i t}{\sum t}$
8. Середній абсолютний приріст	$\overline{\Delta y_i} = \frac{y_i - y_{i-1}}{k}$
9. Середній темп зростання	$\bar{T}(зр) = \sqrt[n-1]{\frac{y_n}{y_1}}$
10. Середній темп приросту	$\bar{T}(пр) = \bar{T}(зр) - 100\%$

Структурно-утворюючі елементи часового ряду

тренд (U_t)

сезонна компонента (S_t)

циклічна компонента (V_t) – коливання відносно тренда з більшою або меншою регулярністю

випадкова компонента (E_t)

Базові компоненти часового ряду $Y_t = f(U_t, S_t, V_t, E_t)$ (5.16)

Врахування додаткових компонент часового ряду $Y_t = f(U_t, S_t, V_t, E_t, Z_t, \eta_t)$

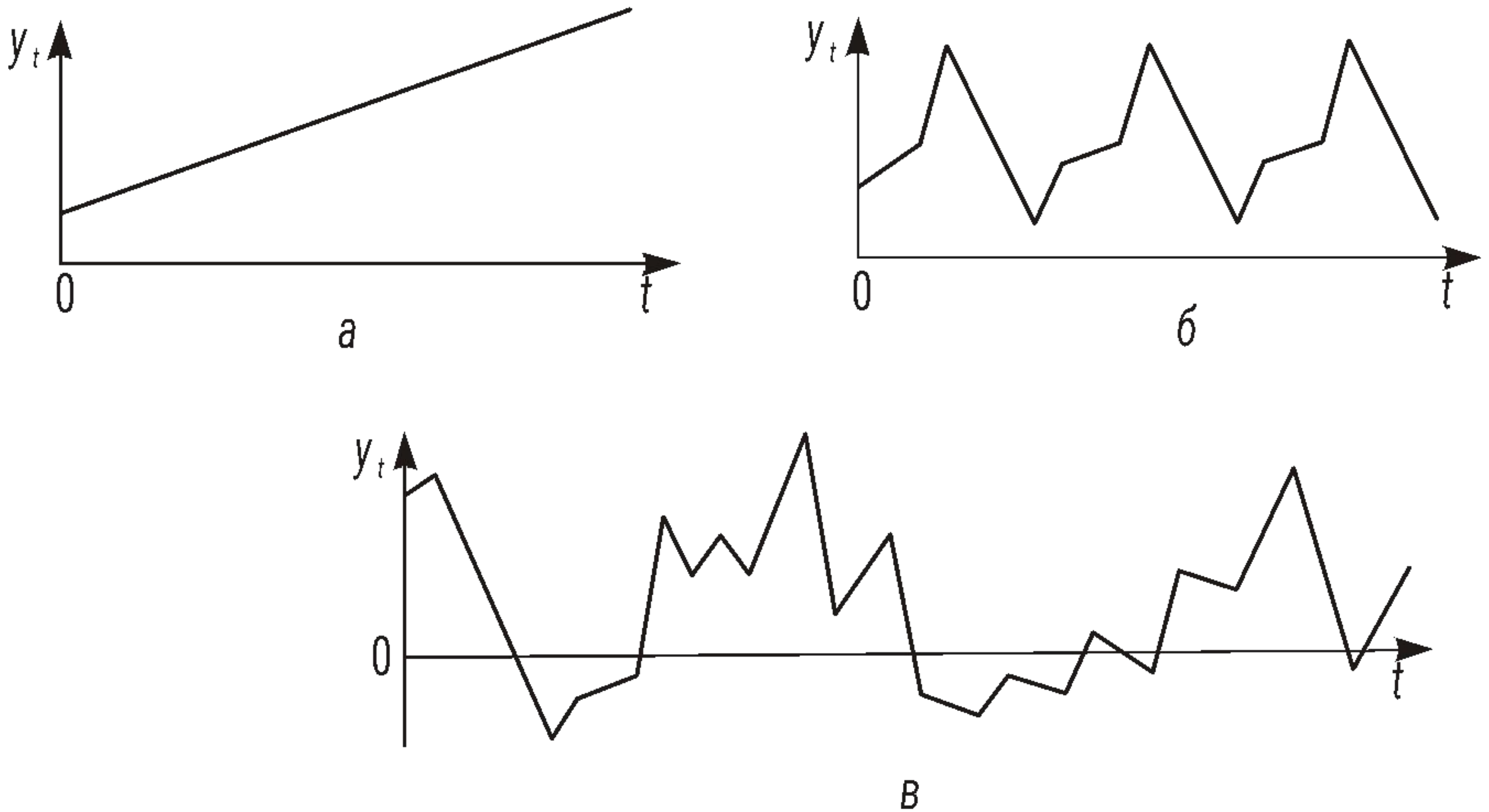
Систематичні та випадкові компоненти часового ряду

$$Y_t = f(U_t, S_t, V_t, E_t)$$

$$Y_t = U_t + S_t + V_t + E_t$$

$$Y_t = U_t + E_t$$

Головні компоненти часового ряду

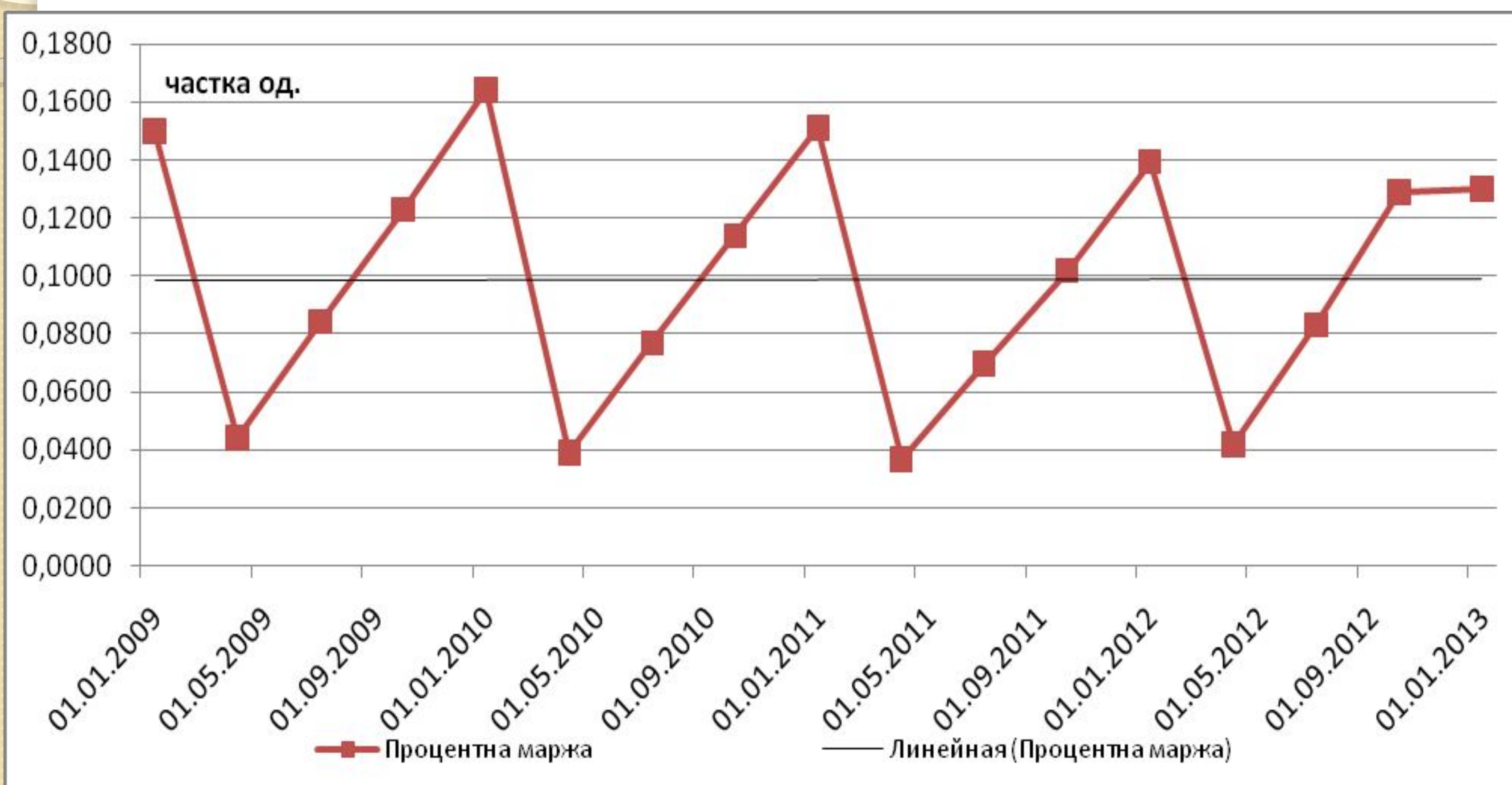


$$Y_t = S_t + E_t$$

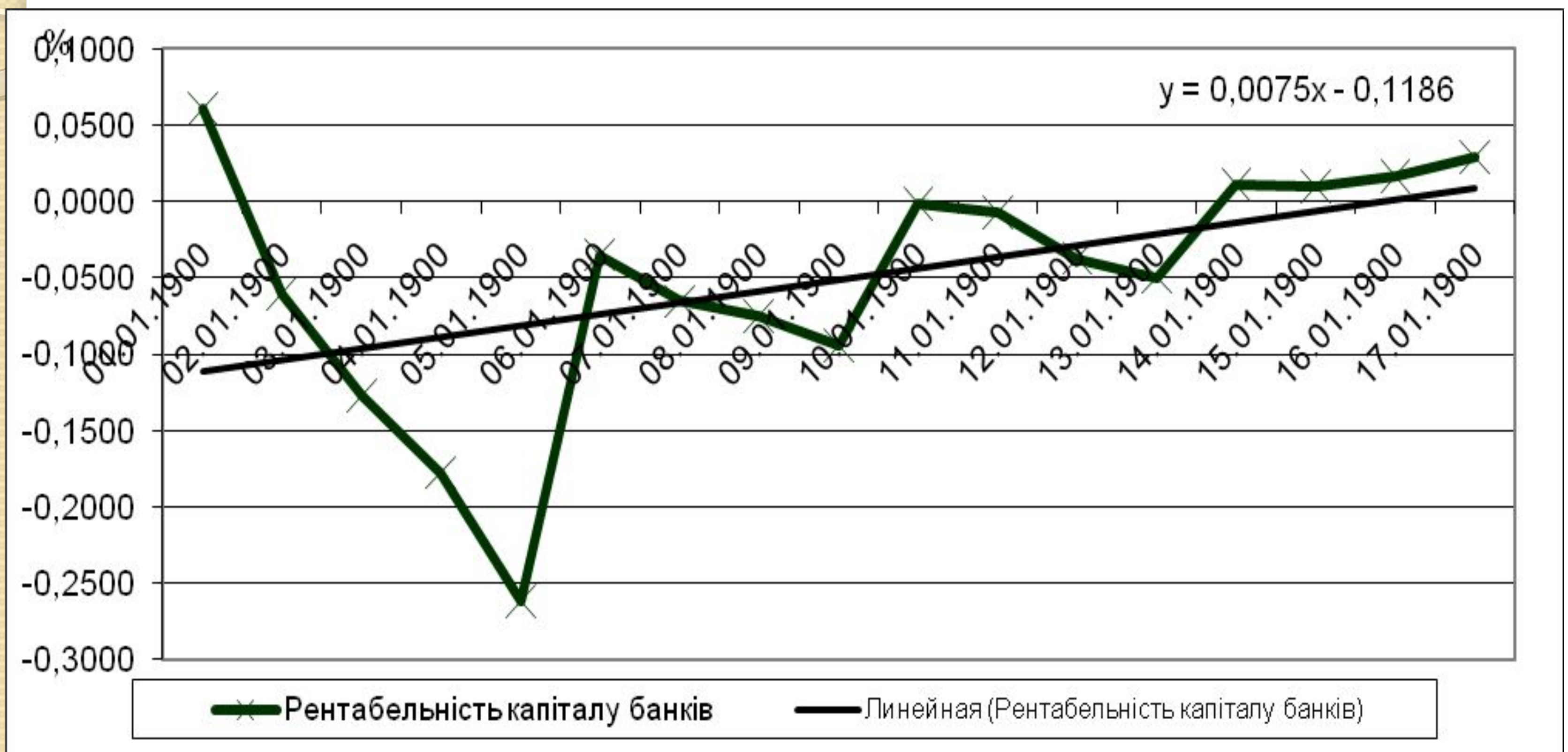
$$Y_t = U_t + S_t + E_t$$

$$Y_t = U_t \cdot S_t \cdot V_t \cdot E_t$$

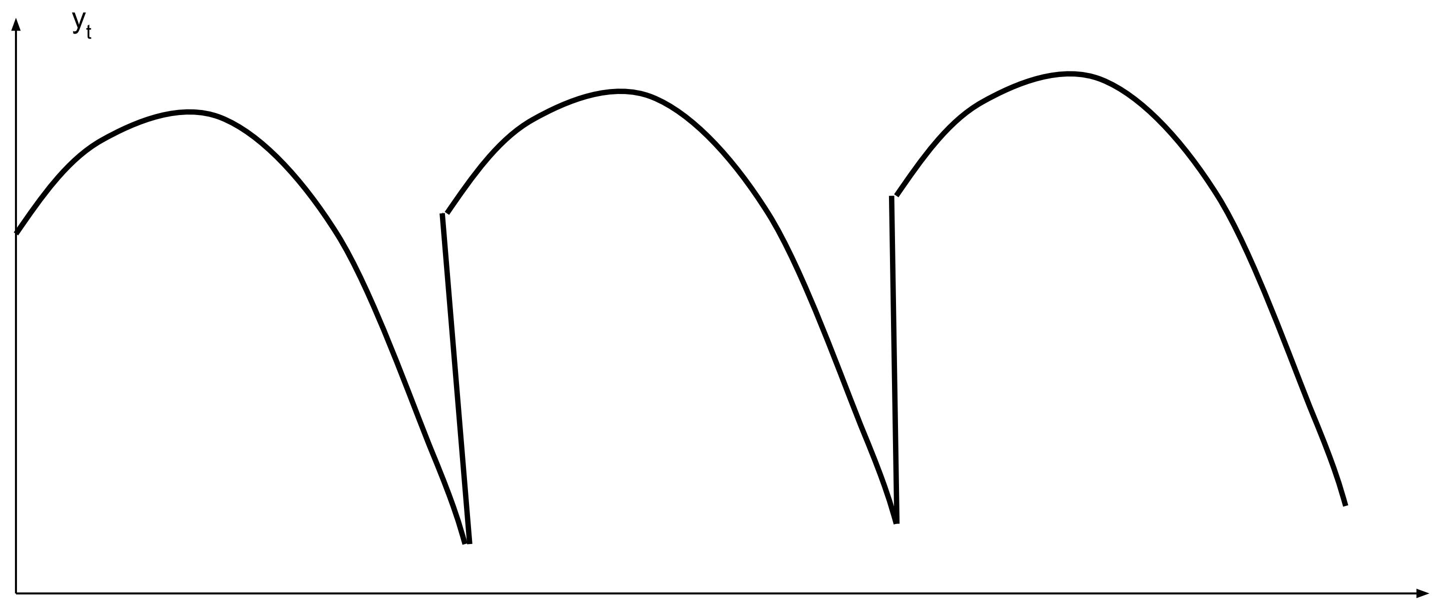
Стаціонарний часовий ряд процентної маржі банків України



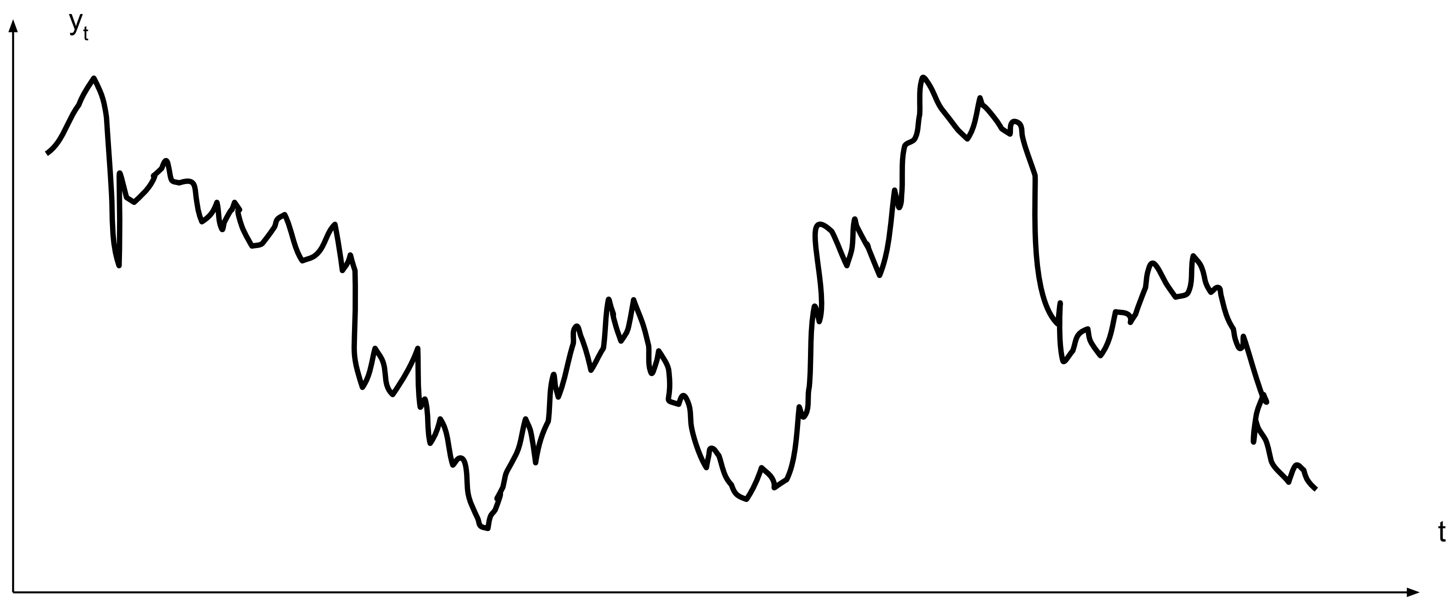
Часовий ряд, що має зростаючу тенденцію

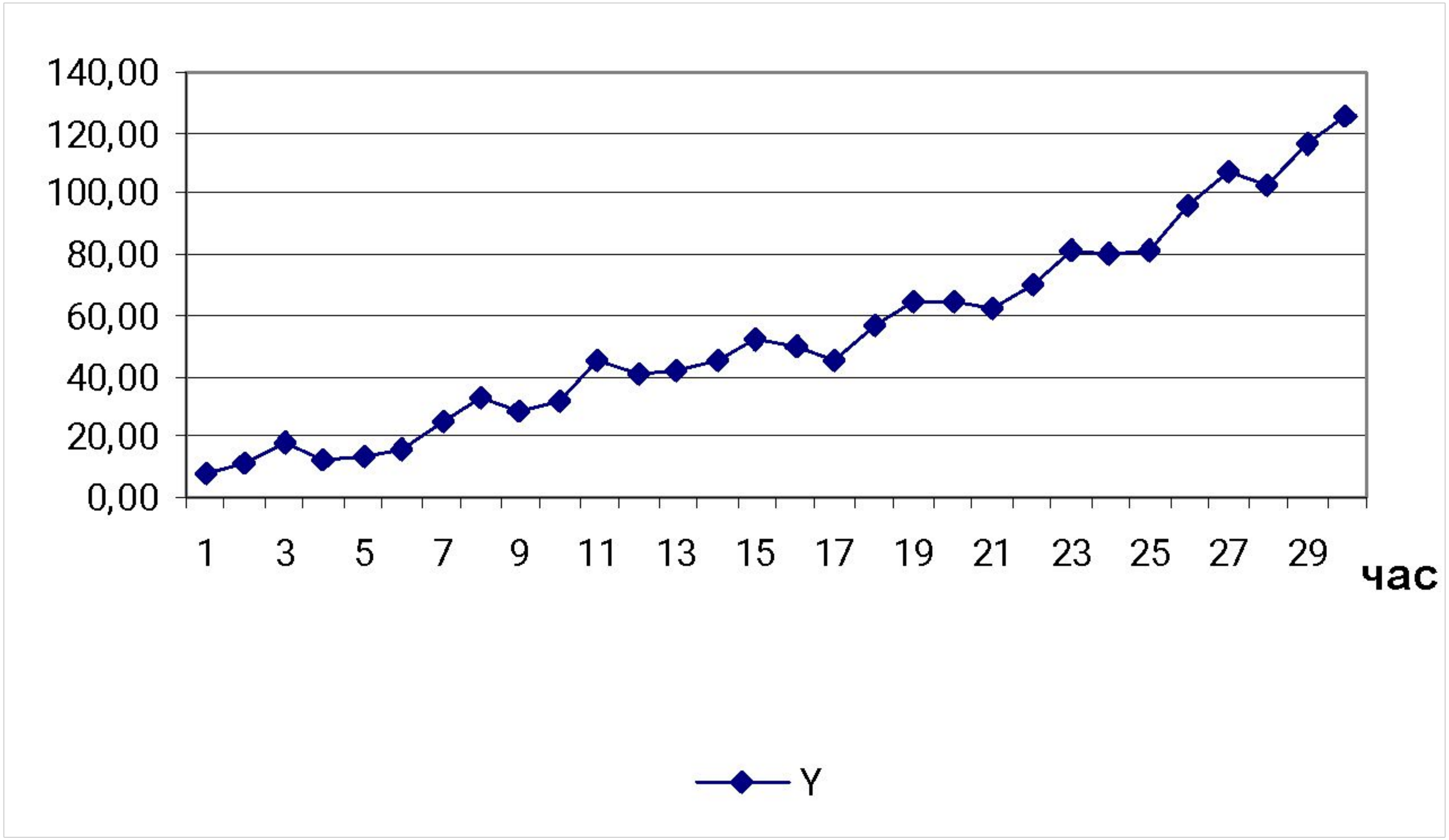


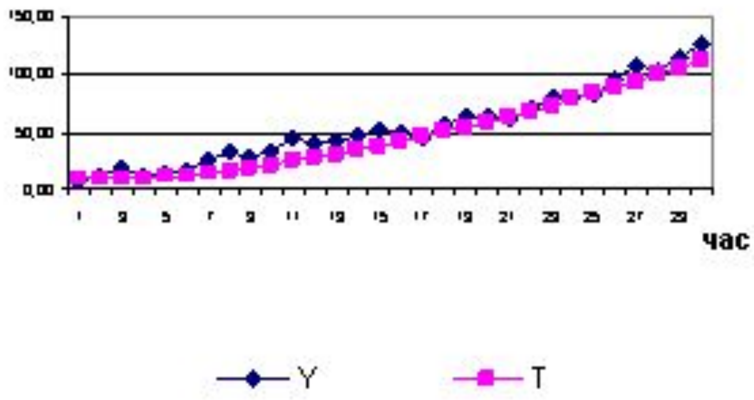
Часовий ряд, що містить тільки сезонну компоненту



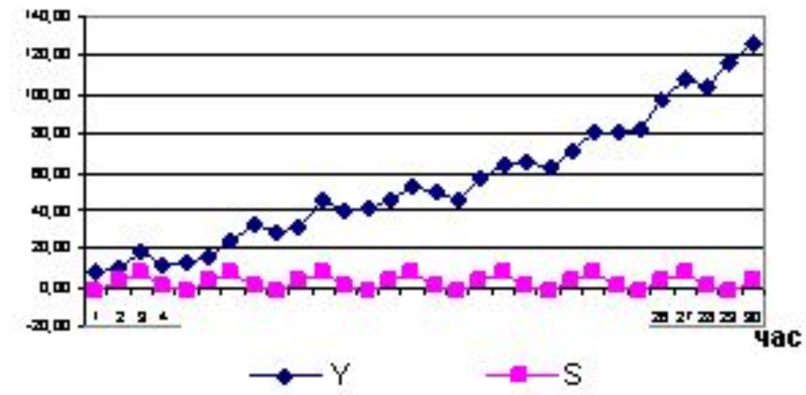
Часовий ряд, що містить випадкову компоненту



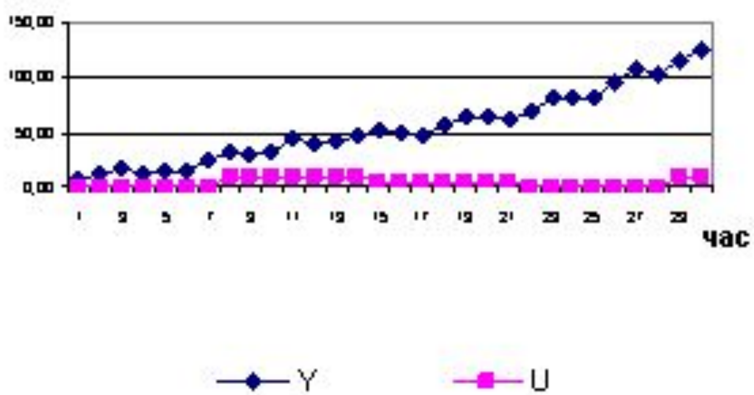

140,00
120,00
100,00
80,00
60,00
40,00
20,00
0,00
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
час
Y



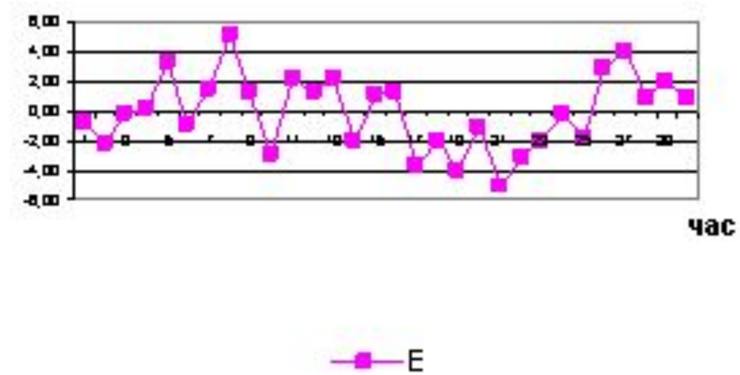
а) трендова компонента $U(t)=3t^{1,3}-5t+12$



б) сезонна компонента



в) циклічна компонента



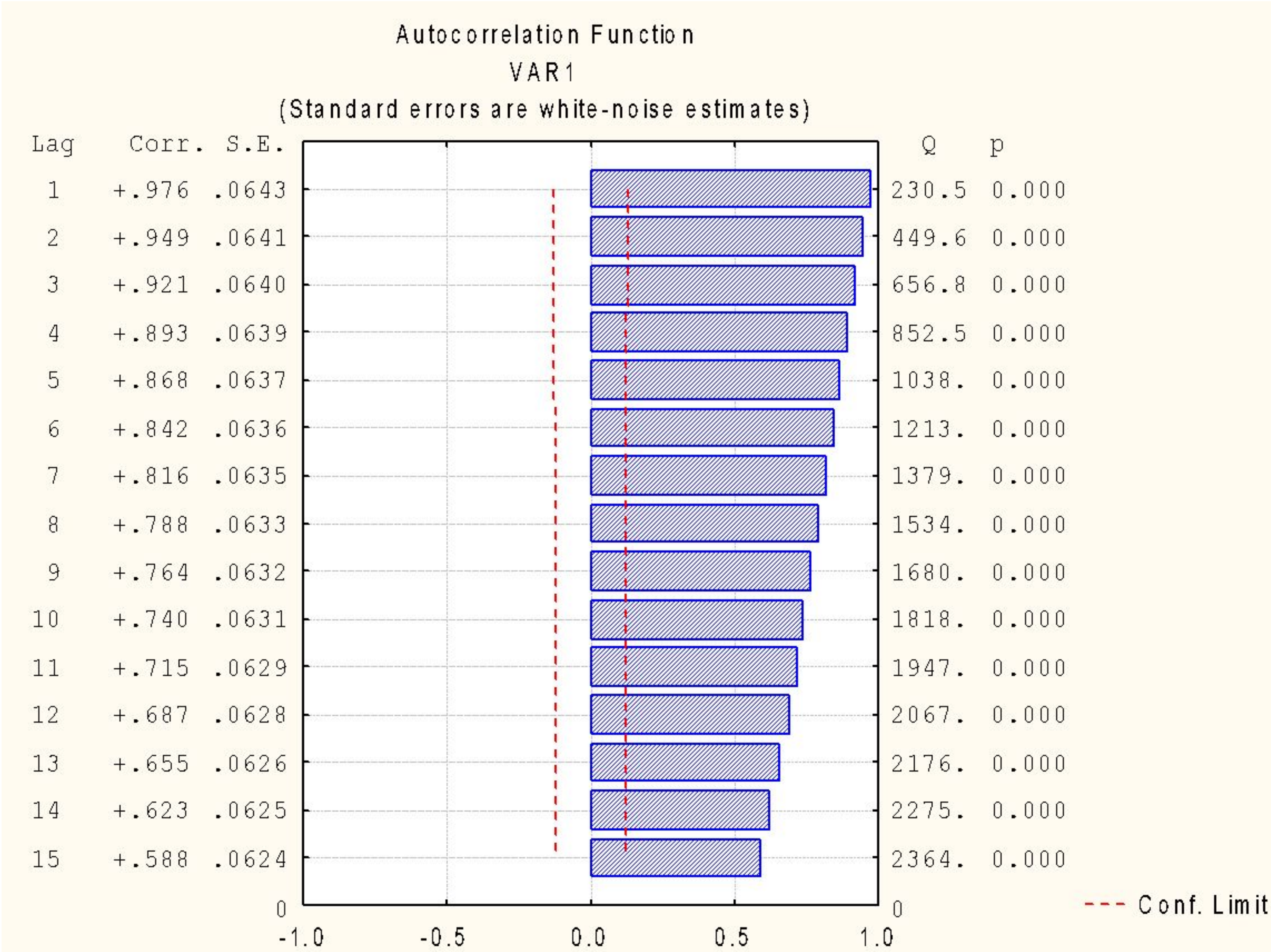
г) випадкова компонента

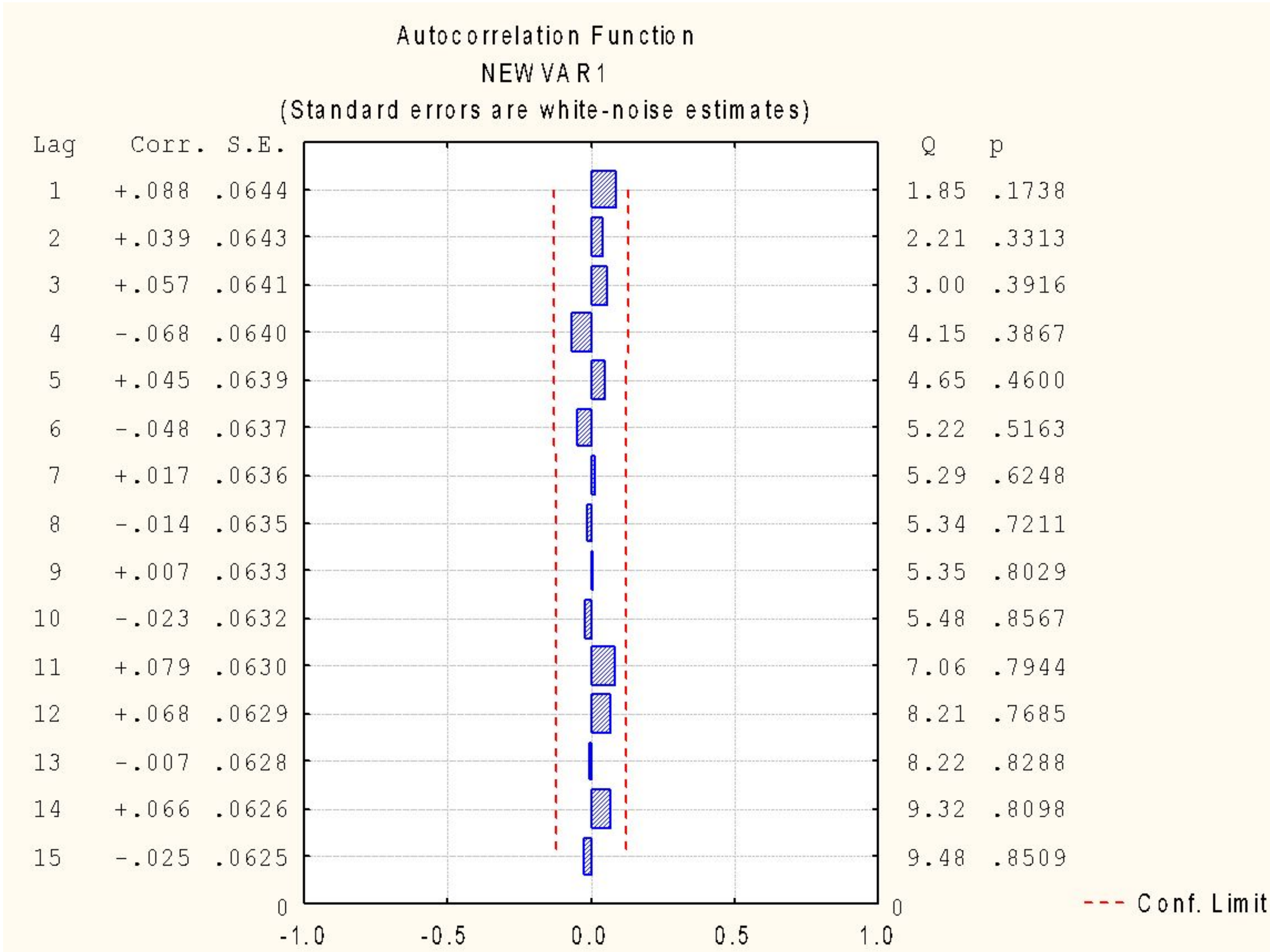
$$r_1 = \frac{\sum_{t=2}^{n} (y_t - \bar{y}_1)(y_{t-1} - \bar{y}_2)}{\sqrt{\sum_{t=2}^{n} (y_t - \bar{y}_1)^2 \sum_{t=2}^{n} (y_{t-1} - \bar{y}_2)^2}},$$

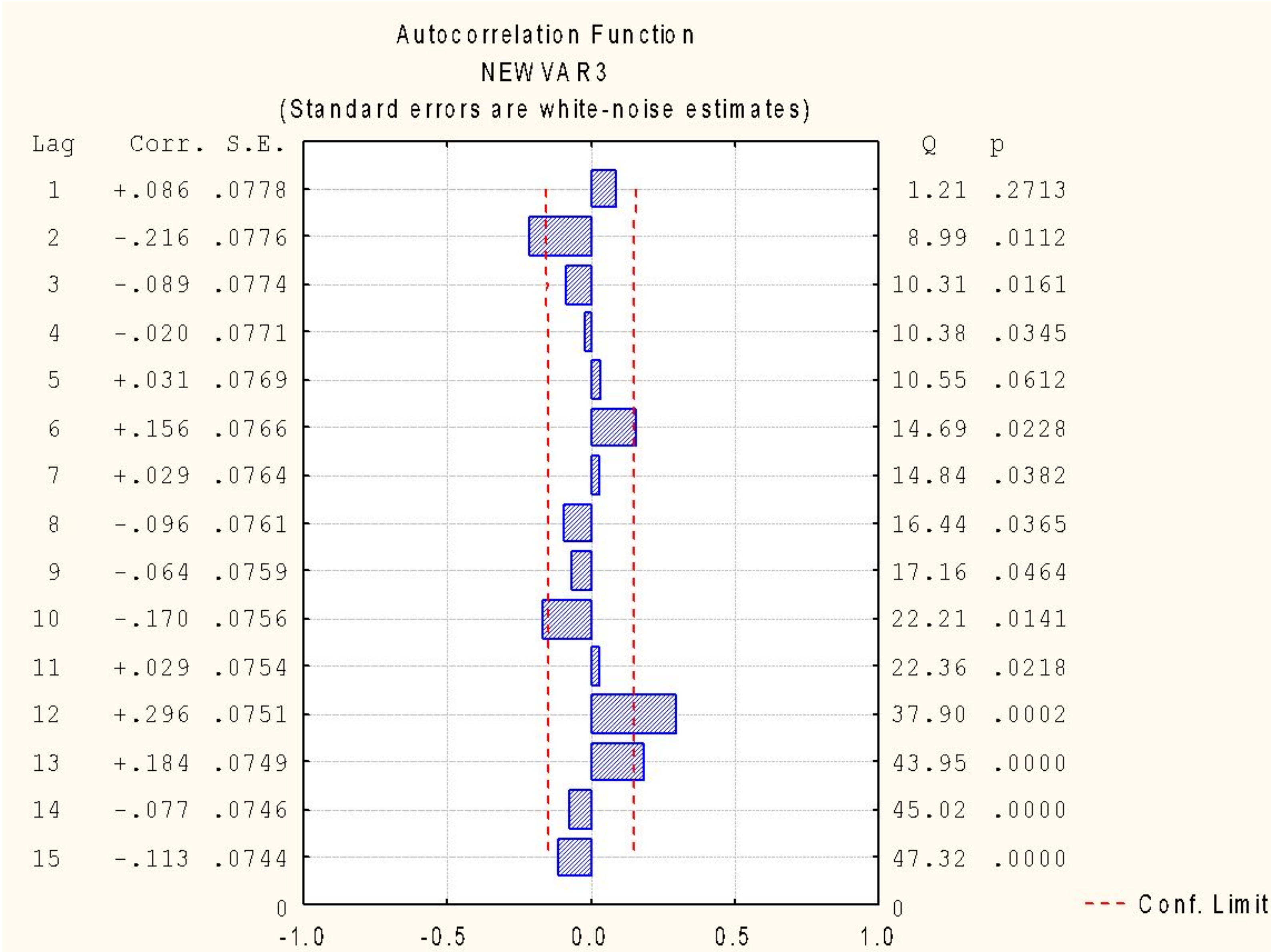
$$\bar{y}_1 = \frac{1}{n-1} \sum_{t=2}^{n} y_t, \qquad \bar{y}_2 = \frac{1}{n-1} \sum_{t=2}^{n} y_{t-1}.$$

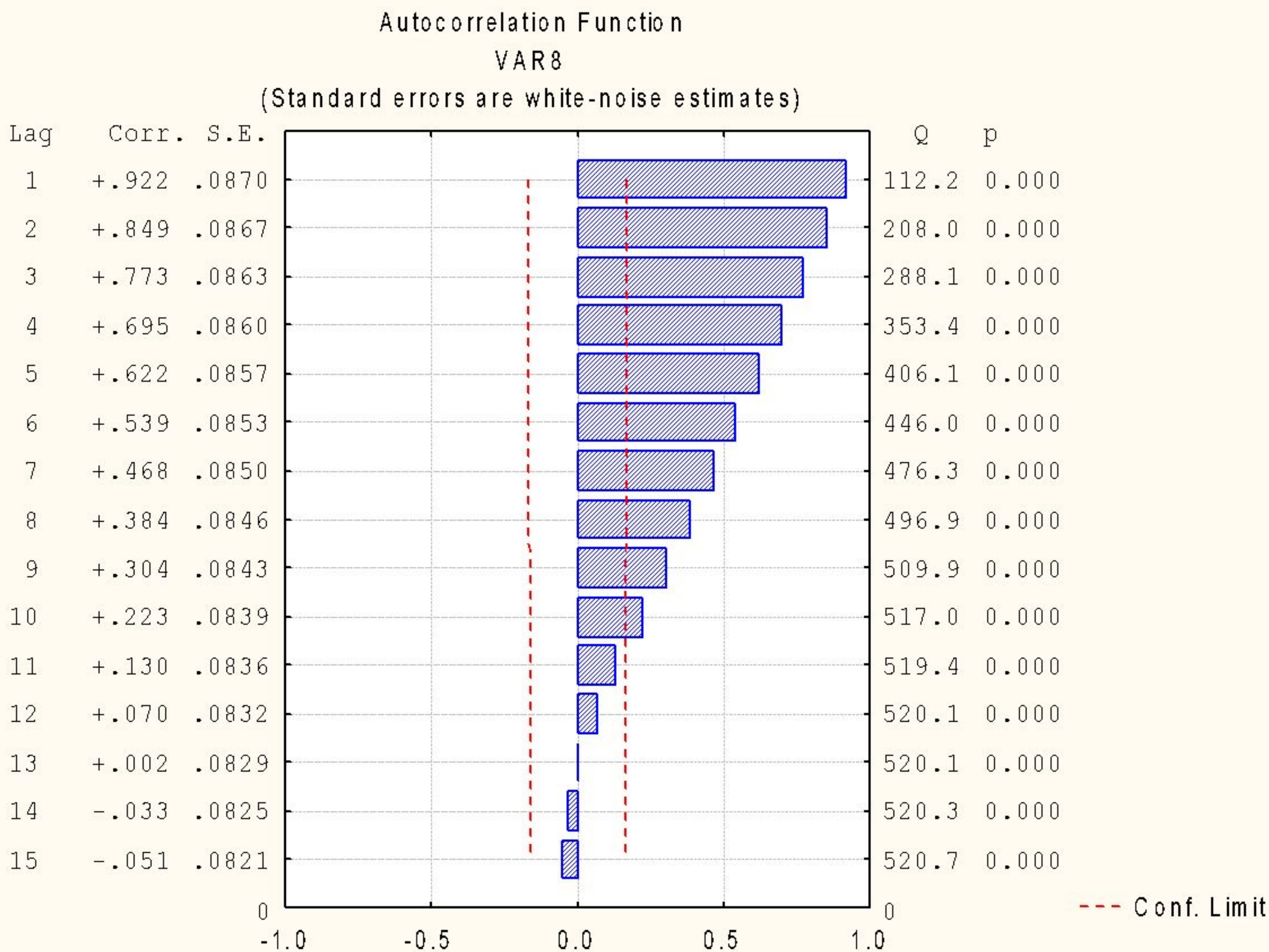
$$r_2 = \frac{\sum_{t=3}^{n}\left(y_t - \overline{y}_3\right)\left(y_{t-2} - \overline{y}_4\right)}{\sqrt{\sum_{t=3}^{n}\left(y_t - \overline{y}_3\right)^2 \sum_{t=3}^{n}\left(y_{t-2} - \overline{y}_4\right)^2}},$$

$$\overline{y}_3 = \frac{1}{n-2}\sum_{t=3}^{n} y_t, \qquad \overline{y}_4 = \frac{1}{n-2}\sum_{t=3}^{n} y_{t-2}.$$


Autocorrelation Function
VAR1
(Standard errors are white-noise estimates)
Lag Corr. S.E. Q p
1 +.976 .0643 230.5 0.000
2 +.949 .0641 449.6 0.000
3 +.921 .0640 656.8 0.000
4 +.893 .0639 852.5 0.000
5 +.868 .0637 1038. 0.000
6 +.842 .0636 1213. 0.000
7 +.816 .0635 1379. 0.000
8 +.788 .0633 1534. 0.000
9 +.764 .0632 1680. 0.000
10 +.740 .0631 1818. 0.000
11 +.715 .0629 1947. 0.000
12 +.687 .0628 2067. 0.000
13 +.655 .0626 2176. 0.000
14 +.623 .0625 2275. 0.000
15 +.588 .0624 2364. 0.000
0
0
-1.0
-0.5
0.0
0.5
1.0
Conf. Limit


Autocorrelation Function
NEWVAR1
(Standard errors are white-noise estimates)
Lag Corr. S.E.
Q p
1 +.088 .0644 1.85 .1738
2 +.039 .0643 2.21 .3313
3 +.057 .0641 3.00 .3916
4 -.068 .0640 4.15 .3867
5 +.045 .0639 4.65 .4600
6 -.048 .0637 5.22 .5163
7 +.017 .0636 5.29 .6248
8 -.014 .0635 5.34 .7211
9 +.007 .0633 5.35 .8029
10 -.023 .0632 5.48 .8567
11 +.079 .0630 7.06 .7944
12 +.068 .0629 8.21 .7685
13 -.007 .0628 8.22 .8288
14 +.066 .0626 9.32 .8098
15 -.025 .0625 9.48 .8509
0
0
-1.0
-0.5
0.0
0.5
1.0
Conf. Limit


Autocorrelation Function
NEWVAR3
(Standard errors are white-noise estimates)
Lag Corr. S.E. Q p
1 +.086 .0778 1.21 .2713
2 -.216 .0776 8.99 .0112
3 -.089 .0774 10.31 .0161
4 -.020 .0771 10.38 .0345
5 +.031 .0769 10.55 .0612
6 +.156 .0766 14.69 .0228
7 +.029 .0764 14.84 .0382
8 -.096 .0761 16.44 .0365
9 -.064 .0759 17.16 .0464
10 -.170 .0756 22.21 .0141
11 +.029 .0754 22.36 .0218
12 +.296 .0751 37.90 .0002
13 +.184 .0749 43.95 .0000
14 -.077 .0746 45.02 .0000
15 -.113 .0744 47.32 .0000
0
0
-1.0
-0.5
0.0
0.5
1.0
Conf. Limit



Autocorrelation Function
VAR8
(Standard errors are white-noise estimates)
Lag
Corr.
S.E.
Q
p
1 +.922 .0870 112.2 0.000
2 +.849 .0867 208.0 0.000
3 +.773 .0863 288.1 0.000
4 +.695 .0860 353.4 0.000
5 +.622 .0857 406.1 0.000
6 +.539 .0853 446.0 0.000
7 +.468 .0850 476.3 0.000
8 +.384 .0846 496.9 0.000
9 +.304 .0843 509.9 0.000
10 +.223 .0839 517.0 0.000
11 +.130 .0836 519.4 0.000
12 +.070 .0832 520.1 0.000
13 +.002 .0829 520.1 0.000
14 -.033 .0825 520.3 0.000
15 -.051 .0821 520.7 0.000
0
0
-1.0
-0.5
0.0
0.5
1.0
Conf. Limit

Методи перевірки стаціонарності часового ряду

Метод перевірки різниць середніх рівнів

$$\bar{y}_1 = \frac{\sum_{t=1}^{n_1} y_t}{n_1} \qquad \hat{\sigma}_1^2 = \sum_{t=1}^{n_1} (y_t - \overline{y_1})^2 / (n_1 - 1)$$

$$\bar{y}_2 = \frac{\sum_{t=1}^{n_2} y_t}{n_2} \qquad \hat{\sigma}_2^2 = \sum_{t=1}^{n_2} (y_t - \overline{y_2})^2 / (n_2 - 1)$$

$$F = \begin{cases} \hat{\sigma}_1^2 \big/ \hat{\sigma}_2^2, & \text{якщо } \hat{\sigma}_1^2 > \hat{\sigma}_2^2 \\ \hat{\sigma}_2^2 \big/ \hat{\sigma}_1^2, & \text{якщо } \hat{\sigma}_1^2 < \hat{\sigma}_2^2 \end{cases} \qquad t = \frac{\left|\bar{y}_1 - \bar{y}_2\right|}{\hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Метод Форстера-Стьюарта

$$k_t = \begin{cases} 1, & \text{якщо } y_t \text{ більше всіх попередніх рівнів} \\ 0, & \text{в іншому разі} \end{cases}$$

$$l_t = \begin{cases} 1, & \text{якщо } y_t \text{ менше всіх попередніх рівнів} \\ 0, & \text{в іншому разі} \end{cases}$$

$$c = \sum_{t=2}^{n} (k_t + l_t) \qquad d = \sum_{t=2}^{n} (k_t - l_t)$$

$$t_c = \frac{\left| c - \hat{\mu} \right|}{\hat{\sigma}_1} \qquad t_d = \frac{\left| d - 0 \right|}{\hat{\sigma}_2}$$

Метод Ірвіна

$$\lambda_t = \frac{|y_t - y_{t-1}|}{\hat{\sigma}_y}; t = 2, 3, \ldots, n,$$

$$\hat{\sigma}_y = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \bar{y})^2}{n}}$$

$$\bar{y} = \frac{\sum_{t=1}^{n} y_t}{n}$$

Модифікований метод Ірвіна

$$\lambda_t = \frac{\left| y_t - y_{t-1} \right|}{\hat{\sigma}_y}$$

$$\overline{y}_t = \frac{(y_{t-1} + y_{t+1})}{2} \Big/ 2 \qquad t = 2,\ 3,\ldots n-1,$$

$$\hat{\sigma}_y = \sqrt{\frac{\sum_{t=1}^{n} (y_{t-1} - \overline{y_t})^2 + (y_{t+1} - \overline{y}_t)^2}{2}}$$

Нормалізація показників часового ряду

Метод нормалізації	Характеристика методу	Математичний опис
1. Зміна інгредієнта	використовується у разі необхідності зміни внутрішньої ознаки інформації – її інгредієнта – на протилежну.	$(-P_{qi})$ $\frac{1}{P_{qi}}$
2. Відносна нормалізація	приведення інформації до безрозмірної нормованої форми базується на понятті «ідеальна» якість, яку можна задавати у вигляді вектора	$P_{qi}^{H} = \frac{P_{qi}}{\max\limits_{q}\{P_{qi}\}}$ $P_{qi}^{H} = \frac{\min\limits_{q}\{P_{qi}\}}{P_{qi}}$
3. Природна нормалізація	використовується для нормалізації показників-стимуляторів	$P_{qi}^{H} = \frac{P_{qi} - \min\limits_{q}\{P_{qi}\}}{\max\limits_{q}\{P_{qi}\} - \min\limits_{q}\{P_{qi}\}}$
4. Нормалізація Севіджа	змінює інгредієнт показника на протилежний; використовується для нормалізації показників-дестимуляторів	$P_{qi}^{H} = \frac{\max\limits_{q}\{P_{qi}\} - P_{qi}}{\max\limits_{q}\{P_{qi}\} - \min\limits_{q}\{P_{qi}\}}$