

А. Т. Яровий, Є. М. Страхов

АНАЛІЗ ЧАСОВИХ РЯДІВ

**Навчально-методичний посібник
для студентів математичних
та економічних спеціальностей**

**Освіта України
2019
Одеса**

Зміст

Вступ	5
1 Основні поняття і компоненти часових рядів	6
1.1 Означення часового ряду та його складові	6
1.2 Тести на наявність тенденції	9
1.2.1 Тест серій, заснований на медіані	9
1.2.2 Тест «висхідних» і «спадних» серій	10
1.2.3 Тест квадратів послідовних різниць (тест Аббе)	11
1.2.4 Тест на порівняння середніх рівнів ряду	12
1.2.5 Тест Фостера — Стюарта	13
1.3 Автокореляція рівнів часового ряду	15
2 Методи побудови тренду	18
2.1 Аналітичне моделювання тенденції часового ряду	18
2.2 Вибір виду кривої	21
2.3 Механічне вирівнювання рівнів часового ряду	24
2.3.1 Згладжування часових рядів за допомогою ковзної середньої	24
2.3.2 Метод експоненціального згладжування (метод Р. Брауна)	30
3 Моделювання сезонних коливань	33
3.1 Побудова аналітичного вигляду сезонної компоненти	40
4 Оцінка адекватності і точності трендових моделей	42
4.1 Оцінка адекватності моделі	42
4.1.1 Перевірка випадковості коливань рівнів залишкової величини	42

4.1.2	Перевірка відповідності розподілу випадкової компоненти нормальному розподілу	43
4.1.3	Перевірка рівності математичного очікування випадкової компоненти нулю	43
4.1.4	Перевірка незалежності значень рівнів випадкової компоненти	44
4.2	Оцінка точності моделі	45
5	Прогнозування часових рядів	49
5.1	Найпростіші методи прогнозування	49
5.1.1	Прогнозування у припущенні незмінності значень попередніх рівнів	49
5.1.2	Прогнозування на підставі середнього рівня ряду	50
5.1.3	Прогнозування методом середнього абсолютного приросту	50
5.1.4	Прогнозування методом середнього темпу зросту	52
5.2	Прогнозування динаміки за трендовими моделями	53
5.2.1	Лінійний тренд	54
5.2.2	Многочлен другого ступеня	54
5.2.3	Многочлен третього ступеня	55
5.2.4	Проста експонента	56
5.2.5	Логарифмічна парабола	57
5.2.6	Модифікована експонента	57
5.2.7	Крива Гомперця	58
5.2.8	Логістична крива Перла — Ріда	58
5.3	Адаптивні моделі прогнозування	61
5.3.1	Метод простого експоненціального згладжування	62
5.4	Прогнозування рядів динаміки, що не мають тенденції	67
5.4.1	Оцінка точності прогнозу	70
6	Приклади побудови моделей часового ряду	74
6.1	Алгоритм побудови моделі часового ряду	74
7	Застосування фіктивних змінних для моделювання сезонних коливань	83
7.1	Моделювання тенденції часового ряду при наявності структурних змін	85

ДОДАТКИ	89
А Завдання для самостійної роботи	89
Б Показники динаміки розвитку економічних процесів	93
В Таблиці математичної статистики	95
В.1 q -квантілі стандартного нормального розподілу	95
В.2 Розподіл Пірсона χ^2	96
В.3 F -розподіл Фішера	97
В.4 Розподіл Ст'юдента	99
В.5 Табульовані значення статистичних критеріїв, що використовуються при обробці сукупностей з грубими похибками	100
В.6 Критерій Аббе	104
В.7 Значення середньої μ і стандартних похибок σ_1 і σ_2 для тесту Фосте- ра — Стюарта	105
В.8 Статистика Дарбіна — Уотсона	106
В.9 Критерій Девіда — Хартлі — Пірсона (RS-критерій)	107
В.10 Критерій Фроціні	108
В.11 Стандартний нормальний розподіл	109

Вступ

Однією з основних задач теорій, що описують економічні процеси, є вміння давати прогноз розвитку процесу. Для цього будують економетричні моделі. Але економісту не завжди відомі чинники, що формують модель, або їх значення. Наприклад, якщо нам необхідно дати прогноз вартості гривні відносно долара на декілька тижнів чи місяців наперед, то будується економетрична модель за допомогою чинників, що формують цю вартість. Але цих чинників може бути дуже багато, а головне, що не завжди відомі їх значення. Тим паче, що значення багатьох чинників формується не в економіці України, а поза її межами. Тому краще з якоюсь періодичністю (раз на тиждень чи місяць) фіксувати вартість гривні відносно долара. Таким чином, отримуємо часовий ряд, який і починаємо досліджувати, а потім будемо математичну модель вартості гривні відносно долара. За отриманою моделлю робимо прогноз вартості гривні.

У посібнику розглянуто складові частини часового ряду: тренд, циклічні і сезонні коливання, випадкова величина і методи їх визначення. Розглянуто алгоритми побудови адитивних і мультиплікативних моделей, досліджується адекватність моделі і питання побудови прогнозу.

Теоретичні положення супроводжуються розв'язанням конкретних прикладів.

Основні поняття і компоненти часових рядів

1.1 Означення часового ряду та його складові

Динамічні процеси, що відбуваються в економіці, проявляються досить часто у вигляді послідовно розташованих у хронологічному порядку значень деякого показника.

Отже, ряд спостережень y_1, y_2, \dots, y_n випадкової величини $\xi(t)$, утворених у послідовні моменти часу t_1, t_2, \dots, t_n , називається часовим рядом. Елементи часового ряду називаються рівнями ряду, а кількість рівнів ряду — довжиною ряду.

Часовий ряд є однопараметрична сім'я випадкових величин $\{\xi(t)\}$. А це означає, що закон розподілу ймовірностей цих випадкових величин може залежати від часу t . Принципових відмінностей часового ряду від послідовності спостережень z_1, z_2, \dots, z_n , що утворюють випадкову вибірку, є дві:

- на відміну від елементів випадкової вибірки члени часового ряду не є статистично незалежними;
- члени часового ряду не однаково розподілені, тобто

$$P\{y(t_1) < y\} \neq P\{y(t_2) < y\} \quad \text{при} \quad t_1 \neq t_2.$$

Це означає, що не можна розповсюджувати властивості і правила статистичного аналізу випадкової вибірки на часові ряди. З іншого боку, взаємозалежність членів часового ряду створює свою специфічну базу для побудови прогнозних значень аналізуючих показників спостережень значень y_1, \dots, y_n .

Розглянемо структуру і класифікацію основних чинників, під впливом яких формуються значення елементів часового ряду.

Розглянемо чотири типи таких чинників:

- 1) *Довгочасові*, що формують загальний напрямок у зміні ознаки y . Вони описуються за допомогою не випадкової функції T_t , як правило, монотонної. Цю функцію називають тенденцією або трендом.
- 2) *Сезонні*, що формують періодичні коливання у певний час року. Будемо позначати результат дії сезонних чинників за допомогою не випадкової функції S_t . Так як ця функція періодична (з періодом кратним «сезонам»), то в її аналітичному виразі беруть участь гармоніки (тригонометричні функції), періодичність яких залежить від змістової суті задачі.
- 3) *Циклічні*, що формують зміну ознаки і обумовлені дією довгочасних циклів економічної, демографічної або астрофізичної природи (хвилі Кондрат'єва, демографічні «ями», цикли сонячної активності). Результат дії циклічних чинників позначимо за допомогою не випадкової функції V_t .

Тенденція, сезонна і циклічна компоненти називаються *регулярними (невипадковими) компонентами* часового ряду.

- 4) Складова частина часового ряду, що залишається після видалення з нього регулярних компонент, є *випадковою величиною, нерегулярною компонентою*, що не піддається обліку. Її дія на формування значень часового ряду якраз і обумовлює стохастичну природу елементів y_t . Результат впливу випадкових чинників позначимо за допомогою випадкових величин ε_t .

Випадкові чинники можуть бути подвійної природи: *несподіваними*, що призводять до стрибкоподібних структурних змін у механізмі формування значень y_t (що виражається в радикальних стрибкоподібних змінах основних структурних характеристик функцій T_t , S_t , V_t розглянутого часового ряду у випадковий момент часу), і *еволюційними*, що обумовлюють відносно невеликі випадкові відхилення значень y_t від тих, що повинні були отриматись тільки під дією чинників T_t , S_t і V_t . Ми будемо розглядати тільки схеми формування часових рядів, що включають в собі дію тільки еволюційних залишкових випадкових чинників.

Якщо регулярні компоненти часового ряду визначені правильно, то після видалення цих компонент з часового ряду залишаються тільки випадкові величини, які матимуть такі властивості:

- випадковість коливань рівнів;

- відповідність нормальному закону розподілу;
- математичне очікування дорівнює нулю;
- незалежність значень рівнів, тобто відсутність автокореляції.

Перевірка адекватності трендових моделей ґрунтується на перевірці виконання у послідовності випадкових величин розглянутих чотирьох властивостей. Якщо хоча б одна з властивостей не виконується, то модель вважається неадекватною, в іншому випадку — адекватною.

Зазначимо, що зовсім не обов'язково, щоб у процесі формування значень часового ряду приймали участь одночасно чинники всіх чотирьох типів. Але у всіх випадках обов'язкова участь випадкових чинників ε_t .

Існує *адитивна* структурна схема впливу чинників на формування значень y_t , яка представляє значення членів часового ряду у вигляді розкладу

$$y_t = T_t + S_t + V_t + \varepsilon_t, \quad t = \overline{1, n}.$$

Також існує *мультиплікативна* схема:

$$y_t = T_t \cdot S_t \cdot V_t \cdot \varepsilon_t, \quad t = \overline{1, n}.$$

Якщо перейти до логарифмів, то мультиплікативну модель можна звести до адитивної:

$$\ln y_t = \ln T_t + \ln S_t + \ln V_t + \ln \varepsilon_t.$$

У цьому випадку не можна допустити, щоб ε_t приймало від'ємні значення. Якщо залишок розглядати як випадкову змінну з нульовим математичним очікуванням, то вид початкової моделі повинен бути таким:

$$y_t = T_t \cdot S_t \cdot V_t \cdot e^{\varepsilon_t}.$$

Ще існує *змішана* модель

$$y_t = T_t \cdot S_t \cdot V_t + \varepsilon_t.$$

Часові ряди поділяються на *стаціонарні* і *нестационарні*.

Випадкові процеси, що протікають в часі приблизно однорідно і мають вигляд випадкових коливань навколо середнього значення без зміни основних характеристик коливань, називаються стаціонарними.

Якщо часовий ряд має явну тенденцію розвитку, то його називають нестационарним.

1.2 Тести на наявність тенденції

Перед тим, як досліджувати часовий ряд, необхідно в'яснити, чи існує у ряді тенденція. Для цього розглянемо декілька тестів.

1.2.1 Тест серій, заснований на медіані

Розглянемо часовий ряд y_1, y_2, \dots, y_n . Створюємо зростаючий варіаційний ряд y_1, y_2, \dots, y_n . Визначимо вибірккову медіану y_{med} за формулою

$$y_{\text{med}} = \begin{cases} y_{\frac{n+1}{2}}, & \text{якщо } n \text{ — непарне,} \\ \frac{1}{2} (y_{\frac{n}{2}} + y_{\frac{n}{2}+1}), & \text{якщо } n \text{ — парне.} \end{cases}$$

Повертаючись до початкового часового ряду, будемо замість кожного його члена y_t ($t = \overline{1, n}$) ставити плюс (+), якщо $y_t > y_{\text{med}}$, і мінус (−), якщо $y_t < y_{\text{med}}$ (члени часового ряду, що дорівнюють y_{med} , в отриманій послідовності плюсів і мінусів не враховуються).

За утвореною послідовністю плюсів і мінусів розраховуємо кількість серій $\nu(n)$ і довжину найдовшої серії $\tau(n)$. Під серією розуміємо послідовність плюсів або мінусів, що йдуть підряд (серія може складатися тільки з одного плюса або тільки з одного мінуса). Зрозуміло, що якщо послідовність складається зі статистично незалежних спостережень, які випадково змінюються біля деякого постійного рівня (тобто тенденція відсутня), то чергування плюсів і мінусів у послідовності повинне бути більш-менш випадковим, тобто ця послідовність не повинна мати досить довгих серій плюсів або мінусів і, відповідно, загальна кількість серій $\nu(n)$ не повинна бути дуже малою.

Сформулюємо алгоритм тесту на присутність тенденції у часовому ряді: якщо хоча б одна з нерівностей

$$\begin{aligned} \nu(n) &> \left\lceil \frac{1}{2} (n + 2 - 1.96\sqrt{n-1}) \right\rceil, \\ \tau(n) &< [1.43 \ln(n+1)], \end{aligned}$$

де $[\cdot]$ — ціла частина числа, не виконується, то з ймовірністю $(1 - \alpha)100\%$, $\alpha \in$

$[0,05; 0,0975]$ у часовому ряді існує тенденція.

Приклад 1.1. Розглянемо часовий ряд:

5; 8; 6; 7; 7; 10; 13; 9; 8; 6; 1; 2; 4; 5; 10; 17; 9; 11; 8; 20; 16.

За допомогою тесту серій перевіримо ряд на наявність тенденції.

Будуємо зростаючий варіаційний ряд:

1; 2; 4; 5; 5; 6; 6; 7; 7; 8; 8; 8; 9; 9; 10; 10; 11; 13; 16; 17; 20.

Так як $n = 21$ — непарне, то $\frac{n+1}{2} = \frac{21+1}{2} = 11$ і це означає, що 11-й член варіаційного ряду і є $y_{\text{med}} = 8$. Повертаємося до початкового часового ряду і згідно з алгоритмом тесту отримаємо таку послідовність (+) і (-):

- - - - + + + - - - - + + + + + .

Маємо $\nu(21) = 4$, $\tau(21) = 6$. Далі,

$$\left[\frac{1}{2} \left(n + 2 - 1.96\sqrt{n-1} \right) \right] = \left[\frac{1}{2} \left(21 + 2 - 1.96\sqrt{21-1} \right) \right] = 7,$$

$$[1.43 \ln(n+1)] = [1.43 \ln(21+1)] = 4.$$

Нерівності мають вигляд:

$$\nu(21) = 4 \not\geq 7,$$

$$\tau(21) = 6 \not\leq 4.$$

Так як нерівності не виконуються, то тенденція у часовому ряді існує. ☑

1.2.2 Тест «висхідних» і «спадних» серій

Цей тест «відчуває» поступове зміщення середнього значення в досліджуваному ряді не тільки монотонного, але і більш загального, наприклад, періодичного характеру.

Розглядається часовий ряд y_1, y_2, \dots, y_n . На i -му місці ставимо плюс (+), якщо $y_{i+1} > y_i$, і мінус (-), якщо $y_{i+1} < y_i$ (якщо два або декілька слідуєчих один за одним члени ряду рівні між собою, то приймається до уваги тільки один з них). Зрозуміло, що послідовність підряд взятих плюсів відповідає збільшенню членів ряду (висхідна серія), а послідовність мінусів — їх зменшенню (спадна серія). Тест

ґрунтується на такому міркуванні: якщо вибірка випадкова (члени ряду незалежні і однаково розподілені), то в утвореній ними послідовності знаків загальна кількість серій не може бути дуже малою, а їх довжина дуже великою.

Тест має такий алгоритм: розрахуємо

$$\nu(n) > \left\lceil \frac{1}{3}(2n - 1) - 1.96\sqrt{\frac{16n - 29}{90}} \right\rceil,$$

$$\tau(n) < \tau_0(n),$$

до $[\cdot]$ — ціла частина числа, $\nu(n)$ і $\tau(n)$ — відповідно загальна кількість серій і кількість підряд взятих плюсів або мінусів у найдовшій серії, а величина $\tau_0(n)$ в залежності від n визначається таким чином:

| | | | |
|-------------|--------------|-------------------|---------------------|
| n | $n \leq 26$ | $26 < n \leq 153$ | $153 < n \leq 1170$ |
| $\tau_0(n)$ | $\tau_0 = 5$ | $\tau_0 = 6$ | $\tau_0 = 7$ |

Якщо хоча б одна з нерівностей не виконується, то часовий ряд має тенденцію з ймовірністю $(1 - \alpha)100\%$, де $0.05 < \alpha < 0.0975$.

Приклад 1.2. Часовий ряд з попереднього прикладу за допомогою тесту «ви-східних» і «спадних» серій перевірити на наявність тенденції:

| | | | | | | | | | | | | | | | | | | | | |
|----|----|----|----|----|-----|-----|----|----|----|----|----|----|----|-----|-----|----|-----|----|-----|----|
| 5, | 8, | 6, | 7, | 7, | 10, | 13, | 9, | 8, | 6, | 1, | 2, | 4, | 5, | 10, | 17, | 9, | 11, | 8, | 20, | 16 |
| + | - | + | | | + | + | - | - | - | - | + | + | + | + | + | - | + | - | + | - |

Маємо $\nu(21) = 10$ і $\tau(21) = 5$.

Отже, $\nu(21) = 10 \not> 10$, $\tau(21) = 5 \not< 5$. Так як нерівності не виконуються, то у часовому ряді існує тенденція. ☑

1.2.3 Тест квадратів послідовних різниць (тест Аббе)

Якщо є підстави вважати, що випадковий розкид членів часового ряду y_1, y_2, \dots, y_n відносно своїх середніх значень підпорядковується нормальному закону розподілу ймовірностей, то доцільно скористатися тестом Аббе. Цей тест більш потужний, ніж попередній.

Спочатку розраховуємо величини

$$\gamma(n) = \frac{q^2(n)}{s^2(n)}, \quad \text{де} \quad q^2(n) = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (y_{i+1} - y_i)^2,$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$$


Якщо виявиться, що $\gamma(n) \leq \gamma_{\alpha}^{\min}(n)$, то у часовому ряді існує тенденція.

Величину $\gamma_{\alpha}^{\min}(n)$ для $n \leq 60$ знаходимо в [табл. В.11](#). При $n > 60$ користуємося формулою

$$\gamma_{\alpha}^{\min}(n) = 1 + \frac{u_{\alpha}}{\sqrt{n + 0.5(1 + u_{\alpha}^2)}},$$

де u_{α} — α -квантиль нормального розподілу.

Приклад 1.3. Перевіримо часовий ряд з попереднього прикладу на наявність тренду за допомогою тесту Аббе.

Маємо $q^2(21) = 10$, $s^2 = 22.63$, $\gamma(21) = 0.442$, $\alpha = 0.05$, $\gamma_{0.05}^{\min}(21) = 0.6574$. Остаточно, $\gamma(21) = 0.442 < \gamma_{0.05}^{\min}(21) = 0.6574$. У часовому ряді з ймовірністю 95 % існує тенденція. 

1.2.4 Тест на порівняння середніх рівнів ряду

За цим тестом можна перевірити часовий ряд на наявність тренду середнього рівня і дисперсій.

Розбиваємо часовий ряд на дві приблизно рівні частини по кількості членів. Вважається, що кожна з них незалежна сукупність з нормальних розподілом.

Якщо часовий ряд має тенденцію, то середні кожної з сукупностей повинні суттєво різнитися між собою. Якщо відмінність несуттєва, то часовий ряд не має тенденції. Формулюємо гіпотези

$$H_0 : \bar{y}_1 = \bar{y}_2, \quad H_A : \bar{y}_1 \neq \bar{y}_2.$$

Далі розраховуємо t_{CT} :

$$t_{CT} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}},$$

де \bar{y}_1 і \bar{y}_2 — середні рівні відповідних частин ряду, n_1 і n_2 — кількість рівнів кожної з частин, σ_1^2 і σ_2^2 — дисперсії першої і другої частин.

Розрахункове значення $t_{СТ}$ порівнюється з $t_{кр} = t(\alpha, n - 2)$, тут n — довжина часового ряду. Якщо $t_{СТ} > t_{кр}$, то з ймовірністю $(1 - \alpha)100\%$ гіпотеза H_0 відхиляється і приймається гіпотеза про існування тенденції середнього рівня.

Для перевірки наявності тенденції у дисперсії формулюємо гіпотези:

$$H_0 : \sigma_1^2 = \sigma_2^2, \quad H_A : \sigma_1^2 \neq \sigma_2^2.$$

Далі розраховуємо статистики: $F_{СТ} = \frac{\sigma_2^2}{\sigma_1^2}$, якщо $\sigma_2^2 > \sigma_1^2$ і $F_{СТ} = \frac{\sigma_1^2}{\sigma_2^2}$, якщо $\sigma_1^2 > \sigma_2^2$. Статистика порівнюється з критичним значенням F -критерію. Якщо $\sigma_2^2 > \sigma_1^2$, то $F_{кр} = F(\alpha; n_2 - 1; n_1 - 1)$, а при $\sigma_1^2 > \sigma_2^2$ $F_{кр} = F(\alpha; n_1 - 1; n_2 - 1)$.


Гіпотеза про рівність дисперсій двох сукупностей відхиляється, якщо $F_{СТ} > F_{кр}$, а це означає, що дисперсія має тенденцію.

Зазначимо, що розглянутий текст доцільно використовувати у випадку з монотонною тенденцією. Якщо часовий ряд змінює загальний напрямок, то точка повороту тенденції може бути близькою до середини ряду, і тоді середні двох частин будуть близькими і тест може не показувати наявності тенденції.

Приклад 1.4. Розглядаємо часовий ряд попереднього прикладу і розбиваємо його на дві частини:

5; 8; 6; 7; 7; 10; 13; 9; 8; 6 і 1; 2; 4; 5; 10; 17; 9; 11; 8; 20; 16

Маємо $\bar{y}_1 = 7.9$, $n_1 = 10$, $\bar{y}_2 = 9.36$, $n_2 = 11$; $t_{СТ} \approx 0.7$, $t_{кр}(0,05; 19) = 2.093$; $t_{СТ} < t_{кр}$. Тенденція середнього відсутня.

Далі, $F_{СТ} = 7.22$, $F_{кр} = F(0,05; 9; 10) = 3.14$; $F_{СТ} > F_{кр}$. Це означає, що існує тренд дисперсії. 

1.2.5 Тест Фостера — Стюарта

Цей тест дає більш надійний результат в порівнянні з попереднім. Крім тенденції у середньому він дозволяє встановити наявність тенденції дисперсії ряду.

Спочатку порівнюємо кожен рівень ряду, починаючи з другого рівня, з усіма попередніми і розраховуємо дві числові послідовності:

$$k_t = \begin{cases} 1, & \text{якщо } y_t \text{ більше за всі попередні рівні,} \\ 0, & \text{в іншому випадку.} \end{cases}$$

$$l_t = \begin{cases} 1, & \text{якщо } y_t \text{ менше за всі попередні рівні,} \\ 0, & \text{в іншому випадку.} \end{cases}$$

Далі розраховуємо величини s і d :

$$s = \sum_{t=2}^n (k_t + l_t), \quad d = \sum_{t=2}^n (k_t - l_t).$$

Величина s характеризує зміну часового ряду і змінюється від 0 до $(n - 1)$, а d характеризує зміну дисперсії рівнів ряду і змінюється від $-(n - 1)$ до $(n - 1)$.

Наступний етап полягає у перевірці гіпотез:

- 1) відхилення величини s від величини μ — математичного очікування величини s для ряду, в якому рівні розташовані випадковим чином;
- 2) відхилення величини d від нуля.

Для перевірки гіпотез використовуємо статистики

$$t_s = \frac{|s - \mu|}{\sigma_1}, \quad \sigma_1 = \sqrt{2 \ln n - 3.4253},$$

$$t_d = \frac{|d - 0|}{\sigma_2}, \quad \sigma_2 = \sqrt{2 \ln n - 0.8456},$$

де μ — математичне очікування величини s , визначеної для ряду, в якому рівні розташовані випадковим чином; σ_1 — середньоквадратичне відхилення для величини s , σ_2 — середньоквадратичне відхилення для величини d .

Для зручності існують табульовані значення величин μ , σ_1 , σ_2 (табл. В.12).

Розрахункові значення t_s і t_d порівнюється з табличними значеннями t -критерію Ст'юдента з рівнем значущості α .

Якщо $t_s > t(\alpha; n - 1)$, то з ймовірністю $(1 - \alpha)100\%$ існує тенденція у середньому, а при $t_d > t(\alpha; n - 1)$ з тією ж ймовірністю існує тенденція у дисперсії.

Приклад 1.5. Розглядаємо все той же часовий ряд

5; 8; 6; 7; 7; 10; 13; 9; 8; 6; 1; 2; 4; 5; 10; 17; 9; 11; 8; 20; 16.

Розрахуємо параметри k_t і l_t :

$$k_t : 1; 0; 0; 0; 1; 1; 0; 0; 0; 0; 0; 0; 0; 0; 1; 0; 0; 0; 1; 0,$$

$$l_t : 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 1; 0; 0; 0; 0; 0; 0; 0; 0; 0.$$

Тоді $s = 6$, $d = 4$.

Статистики t_s і t_d розраховуємо, використовуючи табл. В.12:

$$t_s = \frac{|s - \mu|}{\sigma_1} = \frac{|6 - 5.195|}{1.677} \approx 0.48, \quad t_{KP} = t(0,05; 20) = 2.086,$$

$$t_d = \frac{|d - 0|}{\sigma_2} = \frac{4}{2.279} = 1.76, \quad t_{KP} = t(0,05; 20) = 2.086.$$

Так як $t_s < t_{KP}$ і $t_d < t_{KP}$, то не існує тенденції у середньому і у дисперсії. ☑

1.3 Автокореляція рівнів часового ряду

Одна з головних відмінностей послідовності спостережень часового ряду від випадкової вибірки полягає в тому, що члени часового ряду є статистично взаємозалежними. Кореляційну залежність між послідовними рівнями часового ряду називають *автокореляцією рівнів ряду*. Вона визначається за допомогою лінійного коефіцієнта кореляції між рівнями цього ж ряду, зсунутими на декілька кроків у часі за формулою

$$r_k = \frac{(n-k) \sum_{t=1}^{n-k} y_t y_{t+k} - \sum_{t=1}^{n-k} y_t \sum_{t=1}^{n-k} y_{t+k}}{\sqrt{\left[(n-k) \sum_{t=1}^{n-k} y_t^2 - \left(\sum_{t=1}^{n-k} y_t \right)^2 \right] \left[(n-k) \sum_{t=1}^{n-k} y_{t+k}^2 - \left(\sum_{t=1}^{n-k} y_{t+k} \right)^2 \right]}}.$$

Порядок коефіцієнта автокореляції визначається параметром k . Цей параметр називають *лагом*. Зі збільшення лагу кількість пар значень, за якими розраховується коефіцієнт автокореляції, зменшується. Вважається, що максимальний лаг можна обирати не більше чверті довжини часового ряду.

Послідовність коефіцієнтів автокореляції називається *кореляційною функцією часового ряду*. Графік залежності її значень від величини лагу (порядку коефіцієнта автокореляції) називають *корелограмою*.

Аналіз автокореляційної функції і корелограми дозволяє виявити структуру ряду, а саме:

- 1) **Випадковий ряд.** Якщо часовий ряд є повністю випадковим, тоді $r_k = 0$ для великих k . Для випадкових часових рядів $\{r_k\} \in N(0, \frac{1}{n})$.
- 2) **Короткострокова кореляція.** Стаціонарні часові ряди, тобто $M(y) = a = \text{const}$, досить часто мають короткострокову кореляцію, що характеризується великим значенням r_1 , а r_k асимптотично наближуються до нуля.

- 3) **Нестационарний ряд.** Якщо часовий ряд має тенденцію, то значення r_k не будуть швидко наближатися до нуля навіть при великих значеннях лагу k .
- 4) **Сезонні коливання.** Якщо часовий ряд має сезонні коливання, то корелограма буде коливатися з тією ж частотою. У цьому випадку для визначення коротко-строкової кореляції необхідно з часового ряду видалити сезонні коливання.
- 5) **Викиди.** Якщо часовий ряд має викиди, то корелограма буде мати різкі стрибки. Особливо вони будуть помітні, якщо лаг k буде дорівнювати відстані між викидами.
- 6) **Чергувальний ряд.** Якщо значення часового ряду має нахил до чергування відносно деякого середнього, то корелограма теж буде чергуватись відносно нуля.
- 7) **Лінійна тенденція.** Якщо найбільшим за модулем є коефіцієнт автокореляції першого порядку, то часовий ряд має лінійну тенденцію.

Якщо жоден з коефіцієнтів автокореляції не є значущим, то можна вважати, що часовий ряд не має тенденції і циклічних коливань або він має досить сильну нелінійну тенденцію.

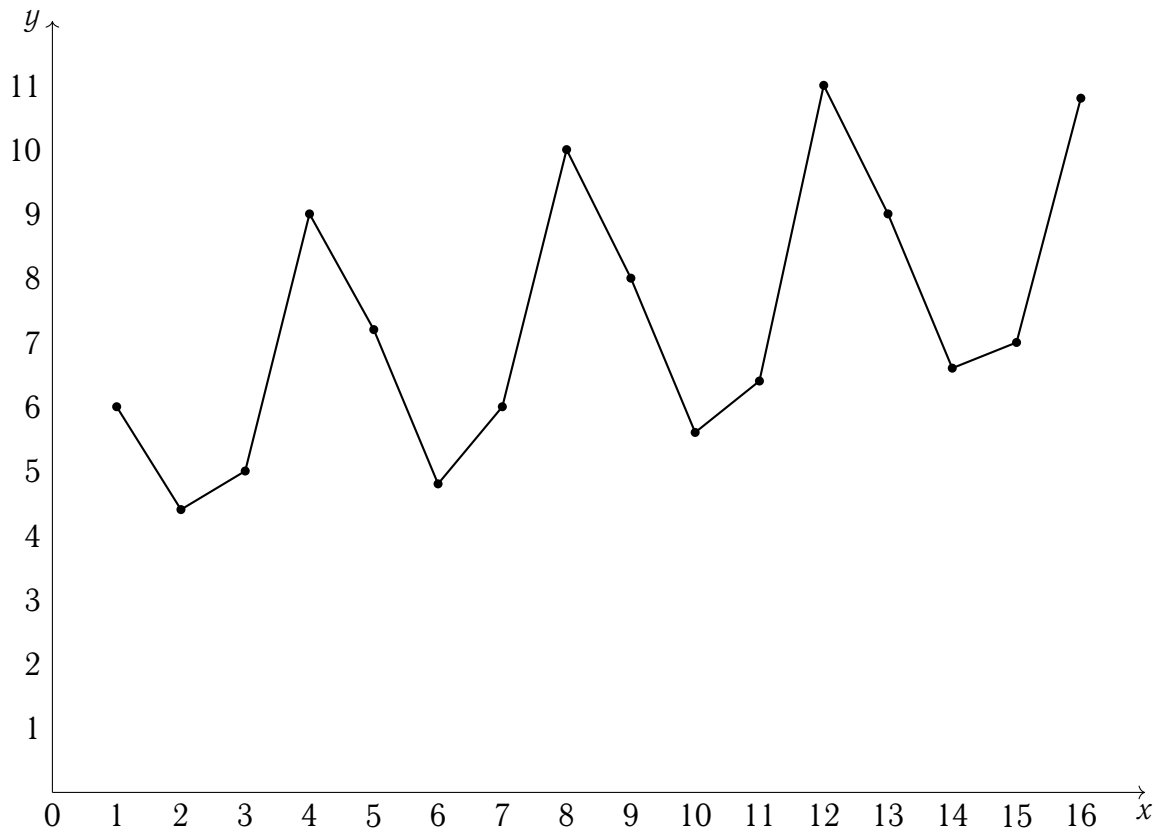
За знаком коефіцієнтів автокореляції не можна робити висновок про зростаючу чи спадну тенденцію в рівнях ряду.

Приклад 1.6. Автокореляційна функція.

Нехай маємо умовні данні про об'єм споживання електроенергії жителями міста за 16 кварталів.

| | y_t | y_{t-1} | y_{t-2} | y_{t-3} | y_{t-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 |

Побудуємо графік функції за цими даними



Розрахуємо коефіцієнти автокореляції першого, другого і т. д. порядків:

| Лаг | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---------------------------------|--------|--------|--------|--------|--------|--------|--------|--------|
| Коефіцієнт автокореляції рівнів | 0.1652 | 0.5669 | 0.1136 | 0.9830 | 0.1187 | 0.7220 | 0.0034 | 0.9738 |

Аналіз значень автокореляційної функції дозволяє зробити висновок про наявність сезонних коливань періодичністю у чотири квартали. ☒

Методи побудови тренду

Досить часто рівні економічних рядів коливаються, і тому тенденція розвитку економічного явища у часі прихована випадковими відхиленнями рівнів в ту чи іншу сторону. З метою більш чітко виявити тенденцію розвитку процесу, а також у подальшому застосування методів прогнозування на основі трендових моделей, проводять згладжування (вирівнювання часових рядів).

Методи виділення не випадкової складової у траєкторії, що відображає поведінку часового ряду можна умовно розбити на дві групи:

- 1) аналітичне вирівнювання з використанням кривої, проведеної між конкретними рівнями ряду так, щоб вона відображала тенденцію, притаманну ряду, і одночасно звільнювала його від незначних коливань;
- 2) механічне вирівнювання окремих рівнів часового ряду з використанням фактичних значень сусідніх рівнів.

2.1 Аналітичне моделювання тенденції часового ряду

Для цього можуть використовуватись такі функції: поліноми різних ступенів, експоненти, логістична крива, крива Гомперця та інші.

Найпростіші поліноміальні криві мають вигляд:

$$\hat{y}_t = a_0 + a_1 t \text{ — поліном першого ступеня,}$$

$\hat{y}_t = a_0 + a_1t + a_2t^2$ — поліном другого ступеня,

$\hat{y}_t = a_0 + a_1t + a_2t^2 + a_3t^3$ — поліном третього ступеня і т. д.

Параметр a_0 — середній рівень ряду динаміки, a_1 називають лінійним приростом, a_2 — прискоренням росту, a_3 — зміна прискорення росту.

Для полінома першого ступеня характерний постійний закон росту, тобто перші прирости (різниці) $u_t = y_t - y_{t-1}$, $t = 2, 3, \dots, n$ є постійна величина і дорівнює a_1 .

Якщо розраховувати перші прирости для полінома другого ступеня, то вони матимуть лінійну залежність від часу, а другі прирости будуть постійними.

Для полінома третього ступеня перші прирости є поліномами другого ступеня, другі прирости є лінійна функція часу, а треті прирости будуть постійні. Характерна особливість поліномів — приріст не залежить від значення функції.

Отже, ми отримали правило вибору ступеня поліному, засноване на величині приростів. Якщо перші прирости (різниці) постійні, то обираємо поліном першого ступеня; поліном другого ступеня, якщо другі прирости постійні і т. д.

Якщо подальший розвиток залежить від значення функції, то використовуються експоненціальні криві. Проста експонента має вигляд

$$\hat{y} = ab^t,$$

де a і b додатні числа. Ордината цієї функції змінюється з постійним темпом приросту:

$$\frac{\hat{y} - \hat{y}_{t-1}}{\hat{y}_t} = 1 - \frac{1}{b}.$$

Логарифм ординати простої експоненти лінійно залежить від часу:

$$\ln \hat{y} = \ln a + t \ln b.$$

Якщо процес характеризується насиченням, то його можна описати кривою, що має асимптоту. Представником сім'ї таких кривих є модифікована експонента. Її рівняння має вигляд

$$y_t = k + ab^t.$$

Ця функція має горизонтальну асимптоту $y = k$. Її графік прямує до асимптоти або при $t \rightarrow \infty$, або при $t \rightarrow -\infty$. Параметр a дорівнює різниці між ординатою кривої (при $t = 0$) і асимптотою. Якщо параметр a від'ємний, то асимптота знаходиться вище кривої, а якщо a додатний, то асимптота проходить нижче її. Параметр b дорівнює відношенню послідовних приростів. Існує чотири варіанти кривої у залежності від

значень b і знаку a . На практиці частіше всього зустрічається випадок, при якому приріст рівня дуже малий, і рівень прямує до деякої границі. Це відбувається якщо $a < 0$ і $b < 1$.

Особливістю модифікованої експоненти є те, що відношення послідовних приростів є постійними, тобто

$$\frac{u_{t_2}}{u_{t_1}} = \frac{u_{t_3}}{u_{t_2}} = \dots = b.$$

Логарифми приростів ординат модифікованої експоненти лінійно залежать від змінної t . Дійсно,

$$u_t = y_t - y_{t-1} = k + ab^t - k - ab^{t-1} = ab^{t-1}(b - 1),$$

звідки

$$\log u_t = \log a + \log(b - 1) + (t - 1) \log b.$$

Більш складним варіантом експоненціальної кривої є логарифмічна парабола:

$$y = ab^t c^{t^2}.$$

Якщо прологарифмуємо, то отримаємо параболу:

$$\log y = \log a + t \log b + t^2 \log c.$$

Розрахуємо темп приросту, отримаємо, що він лінійно залежить від часу t .

Для моделювання часових рядів зі швидким розвитком на початку і затуханням в кінці ряду (тобто вони характеризуються прямуванням до деякої граничної величини), застосовується логістична функція, або функція Перла — Ріда. Вона є зростаючою і записується у вигляді:

$$\hat{y}_t = \frac{k}{1 + ae^{-bt}} \quad \text{або} \quad \hat{y} = \frac{k}{1 + ab^{-t}}, \quad \hat{y} = \frac{k}{1 + 10^{a-bt}}.$$

Тут a і b — додатні параметри, k — граничне значення функції при $t \rightarrow +\infty$. Логістична функція має точку симетрії, що співпадає з точкою перегину: $t = \ln b : a$, $y_t = k : 2$.

У економіці розповсюджені процеси, що спочатку зростають досить повільно, потім прискорюються, а потім уповільнюються. Це, наприклад, процес вводу об'єкта в промислову експлуатацію, процес зміни попиту на товари. Для моделювання таких процесів використовуються S -подібні криві, серед яких виділяється крива Гомперця,

що має такий аналітичний вираз: $\hat{y}_t = ka^{b^t}$, де k , a , b — додатні параметри, $b < 1$, $y = k$ — асимптота кривої. У кривої Гомперця виділяють чотири ділянки: на першій приріст кривої незначний, на другій — приріст збільшується, на третій ділянці приріст майже постійний, а на четвертій — відбувається зменшення темпу приросту і крива необмежено наближається до k .

Логарифм цієї функції є експоненціальна функція, логарифм відношення першого приросту до ординати функції — лінійна функція часу.

Крива Гомперця описує, наприклад, динаміку показників рівня життя, а її модифікації використовують у демографії при моделюванні показників смертності.

Розглянемо декілька застережень про характер процесів, що описуються тими чи іншими типами кривих. Так експоненціальні криві добре описують процеси, що мають «лавиноподібний» характер, а саме, коли приріст залежить в основному уже від досягнутого рівня; при цьому різні обмеження для росту не впливають помітно.

Якщо ж обмежуючий фактор весь час діє і ефективність дії зростає разом з ростом досягнутого рівня, то описувати цей процес можна за допомогою модифікованої експоненти.

А якщо обмежуючий характер починає впливати тільки після деякого моменту (точки перегину), до якого процес розвивався, наближеного до деякого експоненціального закону, то найкраще наближення дають S —подібні криві. Насправді S —подібні криві описують два послідовні лавиноподібні процеси: один з прискоренням розвитку, другий — зі сповільненням.

Модифікована експонента, крива Гомперця і логістична крива при деяких значеннях своїх параметрів мають асимптоти, що проходять вище цих кривих. Такі типи кривих використовуються для опису різного виду процесів «з насиченням».

S —подібні криві використовуються не тільки в роботах, пов'язаних з розвитком популяції, демографічних і страхових розрахунках, але і при розрахунку характеристик різних технічних пристроїв.

2.2 Вибір виду кривої

Розглянемо проблему попереднього вибору виду кривої для часового ряду.

Відбір кривої робиться в два етапи. На першому етапі відбираються функції, що придатні з точки зору змістовного підходу до задачі. На другому етапі для цих функцій розраховується значення критерію і обирається та функція, для якої значення критерію мінімальне.

Частіше за все для обрання функції використовується метод характеристик приросту. Він ґрунтується на порівнянні характеристик зміни приростів часового ряду з відповідними характеристиками обраних функцій.

Вибір виду функції включає попередню статистичну обробку ряду і сам вибір функції. Попередня обробка складається з таких трьох частин: згладжування ряду ковзною середньою, визначення середніх приростів, визначення похідних характеристик приросту.

Згладжування ряду можна робити методом простої ковзної середньої (лінійна апроксимація). Наприклад, для інтервалу згладжування $m = 3$ згладжені рівні розраховуються за формулою

$$\bar{y}_t = \frac{y_{t-1} + y_t + y_{t+1}}{3},$$

а щоб не загубити перший і останній рівні, то їх згладжують за формулою

$$\bar{y}_1 = \frac{5y_1 + 2y_2 - y_3}{6}, \quad \bar{y}_n = \frac{-y_{n-2} + 2y_{n-1} + 5y_n}{6}.$$

Далі розраховуємо перші середні прирости

$$\bar{u}_t^1 = \frac{\bar{y}_{t+1} - \bar{y}_{t-1}}{2}, \quad t = 2, 3, \dots, n-1,$$

другі середні прирости

$$\bar{u}_t^2 = \frac{\bar{u}_{t+1}^1 - \bar{u}_{t-1}^1}{2},$$

а також $\frac{\bar{u}_t^1}{\bar{y}_t}$, $\log \bar{u}_t^1$, $\log \frac{\bar{u}_t^1}{\bar{y}_t}$, $\log \frac{\bar{u}_t^1}{\bar{y}_t^2}$.

Для обрання виду функції скористаємося таблицею.

| Показник | Характер зміни показників у часі | Вид функції |
|--|----------------------------------|--------------------------|
| 1 Перший середній приріст | майже однаковий | поліном першого порядку |
| 2 Перший середній приріст | змінюється лінійно | поліном другого порядку |
| 3 Другий середній приріст | змінюється лінійно | поліном третього порядку |
| 4 $\frac{\bar{u}_t^1}{\bar{y}_t}$ | майже однакові | проста експонента |
| 5 $\log \bar{u}_t^1$ | змінюється лінійно | модифікована експонента |
| 6 $\log \frac{\bar{u}_t^1}{\bar{y}_t}$ | змінюється лінійно | крива Гомперця |
| 7 $\log \frac{\bar{u}_t^1}{\bar{y}_t^2}$ | змінюється лінійно | логістична крива |
| 8 $\frac{\bar{u}_t^1}{\bar{y}_t}$ | змінюється лінійно | логарифмічна парабола |

Для вибору рівняння тенденції, якщо відсутні циклічні коливання, можна скористатися тестом, що ґрунтується на дисперсіях.

Розглянемо загальну варіацію:

$$V_{\text{заг}} = \sum_{t=1}^n (y_t - \bar{y})^2,$$

де \bar{y} — середній рівень часового ряду.

Варіація випадкової величини розраховується за формулою

$$V_{\varepsilon} = \sum_{t=1}^n (y_t - \hat{y}_t)^2,$$

де \hat{y}_t — розрахункові рівні ряду, отримані за рівнянням тенденції. Тоді варіація тренду дорівнює: $V_{\text{тенд}} = V_{\text{заг}} - V_{\varepsilon}$.

Розрахуємо дисперсії:

$$\sigma_{\text{заг}}^2 = \frac{V_{\text{заг}}}{n-1} = \frac{\sum_{t=1}^n (y_t - \bar{y})^2}{n-1}, \quad \sigma_{\varepsilon}^2 = \frac{V_{\varepsilon}}{n-p-1} = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n-p-1},$$

$$\sigma_{\text{тенд}}^2 = \frac{V_{\text{тенд}}}{p-1} = \frac{\sum_{t=1}^n (y_t - \bar{y})^2 - \sum_{t=1}^n (y_t - \hat{y}_t)^2}{p-1},$$

де p — кількість невідомих параметрів обраної кривої.

Далі розраховуємо $F_{СТ}$ і $F_{КР}$:

$$F_{СТ} = \frac{\sigma_{\text{тенд}}^2}{\sigma_{\varepsilon}^2} = \frac{V_{\text{тенд}}(n - p - 1)}{(p - 1)V_{\varepsilon}}, \quad F_{КР} = F(\alpha; p - 1; n - p - 1).$$

Якщо $F_{СТ} > F_{КР}$, то з ймовірністю $(1 - \alpha)100\%$, можна користуватися обраним рівнянням тенденції. При розгляді декількох рівнянь тенденції обирають те, для якого $F_{СТ}$ більша. Або починають аналіз із найбільш простого рівняння тенденції, поступово переходячи до більш складних.

Можна використовувати критерій найменшої суми квадратів відхилень емпіричних рівнів від теоретичних (за обраними кривими):

$$S = \sum_{t=1}^n (y_t - \hat{y}_t)^2 \rightarrow \min.$$

З множини можливих рівнянь тенденції обирають те, якому відповідає найменше значення S .

Зазначимо, що один критерій не забезпечує правильний вибір функції тенденції, тому що існує досить великий вибір функцій для опису тенденції.

Для оцінки параметрів функції тенденції частіше всього використовують метод найменших квадратів.

Відбір найкращого рівняння у випадку нелінійної тенденції можна здійснити шляхом перебору основних форм тенденції і розрахунку для кожного рівняння скоригованого коефіцієнта детермінації за Тейлом \bar{R}_T^2 і відбору рівняння з найбільшим \bar{R}_T^2 .

2.3 Механічне вирівнювання рівнів часового ряду

2.3.1 Згладжування часових рядів за допомогою ковзної середньої

Методи виділення не випадкової складової часового ряду другої групи не пов'язані обмежувальними припущеннями у тому, що загальний вигляд цієї складової ряду відомий досліднику. На «виході» задачі вони дають досліднику тільки алгоритм розрахунку оцінки не випадкової складової у довільній наперед заданій точці t . При згладжуванні цим методом емпіричних даних приймають участь всі члени часового

ряду, відбувається виключення випадкових коливань і визначається значення, що відповідає впливу постійно діючих факторів. Згладжування за допомогою ковзної середньої ґрунтується на простій ідеї: якщо «індивідуальний» розкид значень членів часового ряду y_t біля свого середнього (згладженого) значення характеризується дисперсією σ^2 , то розкид середнього з N членів часового ряду $(y_1 + y_2 + \dots + y_N)/N$ біля того ж середнього буде характеризуватися меншою величиною дисперсії, що дорівнює σ^2/N . А зменшення міри випадкового розкиду (дисперсії) і означає якраз згладження відповідної траєкторії. Тому обирають деяку непарну «довжину усереднення» $N = 2m + 1$, виміряну через підряд ідучі члени часового ряду. А потім згладжені значення часового ряду розраховують по значенням $y_{t-m}, y_{t-m+1}, \dots, y_t, y_{t+1}, \dots, y_{t+m}$ за формулою:

$$\hat{f}_t = \sum_{k=-m}^m \omega_k y_{t+k}, \quad t = m + 1, \dots, n - m, \quad (*)$$

де ω_k ($k = -m, -m + 1, \dots, m$) — деякі додатні «вагові» коефіцієнти, $\omega_k > 0$, $\sum_{k=-m}^m \omega_k = 1$.

При переході від t до $t + 1$ у складі доданків правої частини (*) відбувається заміна тільки одного доданку y_{t-m} доданком y_{t-m+1} , і тому методи, засновані на формулі (*), називають методами ковзної середньої.

Визначення параметрів ω_k засновано на такій процедурі. Відповідно до теореми Вейєрштраса довільна гладка функція при довільних припущеннях може бути локально представлена алгебраїчним поліномом підходящого ступеня p . Тому беремо перші $2m + 1$ членів часового ряду y_1, \dots, y_{2m+1} , будемо за допомогою методу найменших квадратів поліном ступеня p , що апроксимує поведінку цієї початкової частини часового ряду, і використовуємо цей поліном для визначення оцінки \hat{f}_t згладженого значення часового ряду в середній, тобто в $(m + 1)$ -й точці цього відрізка ряду, тобто покладемо $\hat{f}_{m+1} = \hat{y}_{1,m+1}$. Потім «ковзаємо» по осі часу на один такт і таким же способом підбираємо поліном \hat{y}_2 того ж ступеня p до відрізка часового ряду $y_2, y_3, \dots, y_{2m+1}$ і визначаємо оцінку згладженого значення часового ряду в середній точці зсунутого на одиницю відрізка часового ряду, тобто $\hat{f}_{m+2} = \hat{y}_{2,m+2}$ і т. д.

У результаті знайдемо оцінки для згладжених значень \hat{f}_t часового ряду для всіх t , крім $t = 1, 2, \dots, m$ і $t = n, n - 1, \dots, n - m + 1$.

Зазначимо, що значення ω_k ($k = -m, \dots, m$) не залежать від того, для якого саме із «ковзних» часових інтервалів відбувався цей підбір.

Приклад 2.1. Нехай локальна поведінка згладженої функції описується алге-

браїчним поліномом першого ступеня ($p = 1$), тобто на обраному часовому інтервалі довжиною $2m + 1$ часових тактів f_t може бути апроксимована лінійною функцією часу

$$\hat{f}_t = Q_0 + Q_1 t, \quad t = 1, 2, \dots, 2m + 1.$$

Перепозначимо момент часу таким чином, щоб розглядати модель на часовому відрізку $t' = -m, -m + 1, \dots, -1, 0, 1, 2, \dots, m$. Тоді середня точка буде відповідати $t' = 0$. Зрозуміло, що $t' = t - (m + 1)$.

Отже, підберемо коефіцієнти Q_0 і Q_1 таким чином, щоб мінімізувати методом найменших квадратів функцію

$$\sum_{t'=-m}^m [y_{t'} - Q_0 - Q_1 t']^2 \rightarrow \min_{Q_0, Q_1}$$

Диференціюємо ліву частину за Q_0 і Q_1 і прирівнюємо частинні похідні до нуля. Отримаємо систему


$$\begin{cases} (2m + 1)Q_0 + Q_1 \sum_{t'=-m}^m t' = \sum_{t'=-m}^m y_{t'}, \\ Q_0 \sum_{t'=-m}^m t' + Q_1 \sum_{t'=-m}^m (t')^2 = \sum_{t'=-m}^m t' y_{t'}. \end{cases}$$

Так як $\sum_{t'=-m}^m t' = 0$, а оцінка згладженого значення часового ряду визначається в середній точці часового інтервалу відповідним значенням функції $Q_0 + Q_1 t'$ при $t' = 0$, то

$$\hat{f}_{m+1} = Q_0 + Q_1 t' \Big|_{t'=0} = Q_0 = \frac{1}{2m + 1} \sum_{t'=-m}^m y_{t'} = \frac{1}{2m + 1} \sum_{t=1}^{2m+1} y_t.$$

Зробивши ті ж викладки для довільного часового інтервалу $t = k + 1, k + 2, \dots, k + 2m + 1$, ми отримаємо аналогічний результат. Таким чином, у загальному випадку маємо

$$\hat{f}_t = \frac{1}{2m + 1} \sum_{k=-m}^m y_{t+k},$$

тобто при лінійному характері локальної апроксимації часового ряду за значення його згладженого значення в точці t необхідно брати середнє арифметичне з обрамляючих його $2m + 1$ сусідніх значень, $y_{t-m}, \dots, y_t, \dots, y_{t+m}$, тобто $\omega_{-m} = \dots = \omega_m = \frac{1}{2m + 1}$. 

Якщо згладжування проводиться за допомогою лінійної функції, то мова йде про метод простої ковзної середньої.

Одним із найбільш складних питань застосування ковзної середньої для визначення тенденції є вибір інтервалу згладжування. Перед тим, як обрати інтервал згладжування, необхідно зробити економічний аналіз явища, визначити його динаміку, циклічність, величину циклів.

Якщо необхідно згладити мілкі коливання, то інтервал згладжування беруть по можливості великим; інтервал згладжування зменшують, якщо необхідно зберегти більш мілкі коливання. Чим більший інтервал згладжування, тим більш плавним буде тренд.

При виборі інтервалу згладжування необхідно брати до уваги і деякі особливості, властиві методу ковзної середньої. Нехтування ними може привести до спотворення результатів аналізу. Перерахуємо основні особливості.

- 1) Період згладжування повинен бути рівним або кратним періоду коливань, якщо період коливань і його величина постійні.

При непостійному періоді коливань проміжок згладжування необхідно брати середньому періоду коливань. Якщо ця умова не виконується, то згладжений ряд не буде правильно характеризувати загальну тенденцію ряду. У цьому випадку отримаємо не плавний рівень, а новий ряд з тим же періодом, але з більш слабкими коливаннями (мінімумами і максимумами згладженого ряду не завжди будуть співпадати з мінімальними і максимальними значеннями емпіричного ряду).

- 2) При різко мінливих коливаннях проміжок згладжування необхідно збільшити, але не дуже сильно.

Метод ковзної середньої передбачає, що коливання гасять один одного на досліджуваному відрізкові часу. Очевидно, що такий ефект може бути отриманий тільки при простій структурі ряду і його коливань. У хвилеподібних рядах цей метод нерідко приводить до викривлення основної тенденції. Викривлення проявляється в залежності від зміни тенденції ряду. Якщо тенденція ряду має сильний спад або зростання, то значення згладжування ряду перебільшують дійсний рівень ряду динаміки. Якщо ряд не має постійної тенденції сильного спаду або зростання, то спотворення його незначне. Таким чином, метод простої ковзної середньої дає гарні результати в рядах з лінійної тенденцією розвитку.

Зауважимо, що замість середнього можна використовувати медіану значень проміжку згладжування. Основна перевага медіанного згладжування, в порівнянні зі

згладжуванням ковзним середнім, полягає в тому, що результати будуть більш стійкими до викидів на проміжку згладжування. Таким чином, якщо в даних є викиди, то згладжування медіаною дає більш гладкі криві в порівнянні з ковзним середнім. Основний недолік медіанного згладжування в тому, що при відсутності явних викидів, він приводить до більш «зубчатих» кривих і не дозволяє використовувати ваги.

У випадку, якщо тенденція ряду не може бути описана лінійним трендом, використовується зважена ковзна середня.

Для рядів з нелінійною тенденцією необхідно застосовувати метод зважених ковзних середніх. Цей метод має ту перевагу перед методом простих ковзних середніх, що він може дати плавний рівень ряду емпіричних даних з нелінійною тенденцією. Значення зважених ковзних середніх отримують таким чином: всередині кожного інтервалу згладжування часовий ряд вирівнюється поліномом k -ого ступеня

$$\hat{y}_t = \omega_0 + \sum_{i=1}^k \omega_i t^i.$$

Ковзну середню визначають як середній член ряду, утвореного з вирівнених значень. Коефіцієнти поліномів знаходимо з умови

$$\min_{\omega} \sum_{t=1}^n \left(y_t - \sum_{i=1}^k \omega_i t^i \right)^2,$$

тобто методом найменших квадратів.

Приклад 2.2. Розглянемо випадок $p = 2$, $m = 2$. Це означає, що в ковзному часовому відрізкові присутні 5 точок і що локальну поведінку згладженого часового ряду всередині кожного такого відрізка будемо апроксимувати параболою другого порядку:

$$f(t') = \omega_0 + \omega_1 t' + \omega_2 (t')^2, \quad t' = -2, -1, 0, 1, 2.$$

Коефіцієнти ω_i знайдемо з умови:

$$\min_{\omega_0, \omega_1, \omega_2} \left\{ \sum_{t'=-2}^2 [y_{t'} - \omega_0 + \omega_1 t' + \omega_2 (t')^2]^2 \right\}.$$

Маємо

$$\frac{\partial \{\dots\}}{\partial \omega_0} = -2 \left[\sum_{t'=-2}^2 y_{t'} - 5\omega_0 - \omega_1 \sum_{t'=-2}^2 t' - \omega_2 \sum_{t'=-2}^2 (t')^2 \right] = 0,$$

$$\frac{\partial\{\dots\}}{\partial\omega_1} = -2 \left[\sum_{t'=-2}^2 t' y_{t'} - \omega_0 \sum_{t'=-2}^2 t' - \omega_1 \sum_{t'=-2}^2 (t')^2 - \omega_2 \sum_{t'=-2}^2 (t')^3 \right] = 0,$$

$$\frac{\partial\{\dots\}}{\partial\omega_2} = -2 \left[\sum_{t'=-2}^2 (t')^2 y_{t'} - \omega_0 \sum_{t'=-2}^2 (t')^2 - \omega_1 \sum_{t'=-2}^2 (t')^3 - \omega_2 \sum_{t'=-2}^2 (t')^4 \right] = 0.$$

Розв'язок отриманої системи, з урахуванням того, що $\sum_{t'=-2}^2 t' = 0$, $\sum_{t'=-2}^2 (t')^3 = 0$, має вигляд

$$\omega_0 = \frac{1}{35} [-3y_{-2} + 12y_{-1} + 17y_0 + 12y_1 - 3y_2],$$

$$\omega_1 = \frac{1}{10} [-2y_{-2} - y_{-1} + y_1 + 2y_2],$$

$$\omega_2 = \frac{1}{14} [2y_{-2} - y_{-1} - 2y_0 - y_1 + 2y_2].$$

Відповідно, оцінка згладженого значення \hat{f}_t аналізуючого часового ряду в точці t визначається значенням оціненої параболи $\omega_0 + \omega_1 t' + \omega_2 (t')^2$ при $t' = 0$, тобто

$$\omega_0 + \omega_1 t' + \omega_2 (t')^2 \Big|_{t'=0} = \omega_0 = \frac{1}{35} [-3y_{-2} + 12y_{-1} + 17y_0 + 12y_1 - 3y_2].$$

Отримали: $\omega_{-2} = \omega_2 = -\frac{3}{35}$, $\omega_{-1} = \omega_1 = \frac{12}{35}$, $\omega_0 = \frac{17}{35}$.

Отже, можна зробити такий висновок: якщо за $2m + 1$ точками підбирають за допомогою методу найменших квадратів поліном ступеня p , то необхідно мінімізувати критерій виду

$$\sum_{t'=-m}^m [y_{t'} - \omega_0 - \omega_1 t' - \dots - \omega_p (t')^p]^2.$$

Мінімізація приводить до системи, що розбивається на дві більш прості підсистеми через те, що $\sum_{t'=-m}^m (t')^k = 0$ для всіх непарних значень k . Оцінкою невинуватливої складової $\hat{f}(t)$ часового ряду y_t в точці $t = t_0$ буде величина ω_0 . Випишемо значення вагових коефіцієнтів ω_k для деяких значень m і p .

| m | p | ω_{-m} | ω_{-m+1} | \dots | ω_0 |
|-------|---------|----------------|-----------------|--------------------|----------------|
| m_0 | 0 або 1 | $1/(2m_0 + 1)$ | $1/(2m_0 + 1)$ | | $1/(2m_0 + 1)$ |
| 5 | 2 або 3 | $-3/35$ | $12/35$ | | $17/35$ |
| 7 | 2 або 3 | $-2/21$ | $3/21$ | $6/21$ | $7/21$ |
| 9 | 2 або 3 | $-21/231$ | $14/231$ | $39/231$ $54/231$ | $59/231$ |
| 7 | 4 або 5 | $5/231$ | $-30/231$ | $75/231$ | $131/231$ |
| 9 | 4 або 5 | $15/429$ | $-55/429$ | $30/429$ $135/429$ | $179/429$ |



Легко отримати такі властивості ковзних середніх:

- 1) ваги ω_k симетричні відносно середнього значення ω_0 , тобто $\omega_k = \omega_{-k}$, $k = \overline{1, m}$;
- 2) сума усіх ваг дорівнює одиниці: $\sum_{k=-m}^m \omega_k = 1$;
- 3) при одній і тій же довжині $2m + 1$ часового відрізка, за яким робиться усереднення, ваги ω_k для поліномів парного ступеня будуть тими ж самими, що і для поліномів ступеня на одиницю більшого.

Зауваження 2.1 (визначення згладжених значень у крайніх точках). За згладжені значення в m перших і m останніх точках використовуються відповідні значення локально апроксимуючих поліномів, побудованих методом найменших квадратів по $(2m + 1)$ перших і по $(2m + 1)$ останніх точках ряду.

Зауваження 2.2 (усереднення по парній кількості точок). Для розрахунку ковзного середнього ми використовували непарну кількість точок — $(2m + 1)$. Але інколи необхідно взяти парну кількість точок.

У цих випадках згладжуване значення часового ряду, як і раніше, розраховується в середній точці ковзного інтервалу усереднення. Очевидно, що такою (при довжині інтервалу згладжування $2m$ і при початковій точці k цього інтервалу) буде точка $t^* = \frac{k + (k + 2m) - 1}{2} = k + m - \frac{1}{2}$, $k = 1, 2, \dots, n - (2m - 1)$.

Тобто ми отримуємо згладжені значення часового ряду не в точках його спостереження $t = 1, 2, \dots, n$, а для моментів часу t^* , що лежать посередині між точками спостереження. Щоб отримати згладжені значення часового ряду в одній з точок його спостереження необхідно розрахувати його згладжені значення для двох сусідніх точок і взяти їх середнє значення:

$$\hat{f}(k + m) = \frac{1}{2} \left[f \left(k + m - \frac{1}{2} \right) + f \left(k + m + \frac{1}{2} \right) \right].$$

Така процедура називається *центруванням*.

2.3.2 Метод експоненціального згладжування (метод Р. Брауна)

Особливість цього методу полягає в тому, що при згладжуванні використовуються значення тільки попередніх рівнів ряду, взятих з вагами, що зменшуються по мірі віддалення їх від моменту часу, для якого визначається згладжене значення рівня ряду. Якщо для початкового часового ряду y_1, \dots, y_n відповідні згладжені

значення рівнів позначити через S_t , $t = \overline{1, n}$, то експоненціальне згладжування здійснюється за формулою $S_t = \alpha y_t + (1 - \alpha)S_{t-1}$, де $\alpha \in (0, 1]$ — параметр згладжування, а величина $(1 - \alpha)$ називається коефіцієнтом дисконтування.

Маємо, що середній рівень ряду на момент t дорівнює лінійній комбінації двох величин: фактичного рівня для цього ж моменту і середнього рівня, розрахованого для попереднього періоду. Таким чином, середня формується під впливом всіх попередніх рівнів ряду від його початку і до моменту t включно. Тобто, середня для моменту t є лінійною комбінацією значень всіх спостережень від y_1 до y_t . Формула для S_t є рекурентною. Послідовно розкриваючи значення Q_{t-1} , отримаємо:

$$\begin{aligned} Q_t &= \alpha y_t + (1 - \alpha)Q_{t-1} = \alpha y_t + (1 - \alpha) [\alpha y_{t-1} + (1 - \alpha)Q_{t-2}] = \\ &= \alpha y_t + \alpha(1 - \alpha)y_{t-1} + (1 - \alpha)^2 [\alpha y_{t-2} + (1 - \alpha)Q_{t-3}] = \dots = \\ &= \alpha y_t + \alpha(1 - \alpha)y_{t-1} + (1 - \alpha)^2 y_{t-2} + \dots + \alpha(1 - \alpha)^k y_{t-k} + \dots + (1 - \alpha)y_0. \end{aligned}$$

Величина y_0 є величиною, що характеризує початкову умову. Перепишемо S_t в іншому вигляді:

$$Q_t = \alpha \sum_{i=0}^{t-1} (1 - \alpha)^i y_{t-i} + (1 - \alpha)^t y_0.$$

Останній вираз пояснює, чому середня Q_t називається експоненціальною: відносна вага окремого спостереження дорівнює $\alpha(1 - \alpha)^i$ і зменшується при віддаленні спостережень у минуле відповідно експоненціальній функції. При достатньо далекому віддаленні в минуле від моменту t вага буде настільки малою, що відповідний рівень ряду практично не впливатиме на Q_t .

Експоненціальна середня легко адаптується до нових умов (при русі в часі). Запишемо вираз для Q_t у формі

$$Q_t = Q_{t-1} + \alpha (y_t - Q_{t-1}).$$

Це означає, що експоненціальна середня на момент t є експоненціальна середня попереднього періоду плюс поправка. Вона є часткою від різниці нинішнього спостереження і експоненціальної середньої попереднього моменту. Якщо Q_{t-1} розглядати як прогноз на один крок вперед, то $y_t - Q_{t-1}$ є похибка цього прогнозу. Отже, прогноз для часу $(t + 1)$ (дорівнює Q_t) враховує спостережену в момент t похибку.

Можна довести, що математичне очікування величин Q_t і y_t співпадають. Крім того, доведено, що дисперсія експоненціальних середніх менша, ніж дисперсія спо-

стерезень:

$$\sigma^2(Q_t) = \frac{\alpha}{2 - \alpha} \sigma^2(y_t).$$

Зрозуміло, що при великих значеннях α (близьких до 1) дисперсія експоненціальних середніх незначно відрізняється від дисперсії спостережень. Чим менше α , тим більше цей коефіцієнт відіграє роль «фільтра», що поглинає коливання.

Практика показала, що параметр α краще обирати з проміжку $[0.1; 0.3]$. При виборі значення α необхідно враховувати, що для підвищення швидкості реакції на зміну процесу розвитку необхідно збільшити значення α , тобто збільшуємо вагу поточних спостережень, але при цьому зменшуються «фільтраційні» можливості експоненціальної середньої.

Зазначимо, що Р. Браун пропонує обирати α таким чином: $\alpha = \frac{2}{n + 1}$.

Початкове значення y_0 обирається рівним значенню першого рівня ряду y_1 або середньому арифметичному декількох перших рівнів ряду.

Указаний спосіб вибору y_0 забезпечує добру погодженість згладженого і початкового ряду для перших рівнів. Якщо при переході до правого кінця часового ряду згладжені значення при обраному значенні параметра α значно відрізняються від відповідних значень початкового ряду, то необхідно обрати інше значення параметра α .

Ми розглянули формули, що визначають експоненціальні середні першого порядку, тобто середні, що отримуються безпосередньо при згладжуванні початкових спостережень. Взагалі-то доцільно користуватися експоненціальними середніми більш високих порядків. Експоненціальна середня порядку k має вигляд

$$Q_t^{(k)} = \alpha \sum_{i=0}^{\infty} (1 - \alpha)^i Q_{t-1}^{(k-1)},$$

або такий, після перетворення:

$$Q_t^{(k)} = \alpha Q_t^{(k-1)} + (1 - \alpha) Q_{t-1}^{(k)}.$$

Моделювання сезонних коливань

При розгляді квартальних або місячних даних досить часто спостерігаються коливання, що повторюються і на протязі великого проміжку не змінюються. У статистиці ці внутрішньорічні зміни називають сезонними коливаннями.

Вивчення сезонних коливань відбувається за двома напрямками. Перший — аналіз сезонних коливань як задача, що має самостійне значення. Другий напрямок виключення впливу сезонності з ряду динаміки для подальшого вивчення динаміки.

Сезонність характеризується тривалістю періоду коливань, амплітудою сезонних коливань і розташуванням максимумів і мінімумів у часі.

Якщо сезонні коливання в усіх трьох аспектах постійні, то сезонність має постійний характер, а якщо розглянуті показники з часом змінюються, то сезонність змінна. У залежності від виду сезонності застосовують різні статистичні методи для визначення сезонного компонента.

Найбільш розповсюдженими методами виміру сезонних хвиль без попереднього виключення загальної тенденції є такі:

- 1) метод абсолютних різниць;
- 2) метод відношення середніх місячних до середньої за весь період;
- 3) метод відношень щомісячних рівнів до середньої року.

У цих методах вимір сезонних хвиль ґрунтується на застосуванні середньої арифметичної.

Існує ще одна група методів, що вимірює сезонні хвилі, що ґрунтуються на застосуванні відносних величин:

- 1) метод відносних величин;
- 2) метод відносних величин на основі медіани;
- 3) метод У. Персонса.

Ці методи можна застосовувати тільки у випадку, якщо загальна тенденція ряду мало змінюється або має лінійну тенденцію.

Якщо тенденція ряду має складний характер, то у цьому випадку перед розрахунком сезонної хвилі тенденцію ряду елімінують (видаляють).

При виявленні тренду ряду, що має сезонні коливання, можна скористатися методом ковзної середньої або методом найменших квадратів. Однак метод ковзної середньої недосконалий. Цей метод не дозволяє повністю звільнитися від сезонних коливань.

Чим більш правильно розрахуємо тренд, тим більш точно визначимо сезонні коливання.

Крім визначення тренду необхідно ще встановити тип зв'язку компонентів ряду.

У залежності від типу зв'язку компонентів ряду динаміки (адитивного чи мультиплікативного) сезонні коливання виражаються в різних показниках. Якщо зв'язок адитивний, то сезонні коливання розраховуються в абсолютних величинах, а якщо мультиплікативний, то у відносних величинах (у %).

Тип зв'язку можна визначити по динаміці відхилень фактичних рівнів від тренду $(y_t - \hat{y}_t)$ або $\frac{y_t - \hat{y}_t}{y_t}$. Якщо абсолютні відхилення мають тенденцію до зросту, а відносні змінюються приблизно на одному рівні, то це свідчить про мультиплікативний тип зв'язку тренду і сезонної компоненти.

Якщо амплітуда коливань приблизно постійна, то будують адитивну модель часового ряду, в якій значення сезонної компоненти постійне для різних циклів. Але якщо амплітуда сезонних коливань зростає або зменшується, то будують мультиплікативну модель ряду.

Виявлення наявності у часовому ряді сезонних коливань зводиться до перевірки на випадковість залишкового ряду $y_t - T_t$.

Розрахунок сезонної складової можна зробити різними методами. Розглянемо деякі з них.

1. Метод абсолютних різниць.

Розглянемо дані про споживання молока в регіоні за п'ять років.

| місяць | 2013 | 2014 | 2015 | 2016 | 2017 | \bar{y}_t | ланки
сезонної
хвилі | I_s
% | $\frac{y_t}{\bar{y}_{2013}}$ | $\frac{y_t}{\bar{y}_{2014}}$ | $\frac{y_t}{\bar{y}_{2015}}$ | $\frac{y_t}{\bar{y}_{2016}}$ | $\frac{y_t}{\bar{y}_{2017}}$ | \sum
5 років | I_s
% |
|--------------------------------------|--------|--------|--------|--------|--------|-------------|----------------------------|------------|------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|-------------------|------------|
| січень | 117.8 | 125.1 | 126.8 | 131.8 | 138.7 | 128.04 | -16.01 | 88.9 | 86.6 | 88.8 | 88.8 | 89.3 | 90.7 | 444.2 | 88.8 |
| лютий | 108.1 | 118.4 | 118.2 | 121.3 | 129.8 | 119.16 | -24.89 | 82.7 | 79.5 | 84.1 | 82.8 | 82.2 | 84.8 | 413.4 | 82.7 |
| березень | 127.9 | 136.5 | 138.6 | 142.6 | 149.7 | 139.06 | -4.99 | 96.5 | 94.0 | 96.9 | 97.1 | 96.6 | 97.8 | 482.4 | 96.5 |
| квітень | 132.4 | 138.4 | 139.8 | 144.6 | 150.8 | 141.20 | -2.85 | 98.0 | 97.4 | 98.3 | 97.9 | 98.0 | 98.6 | 490.2 | 98.0 |
| травень | 162.8 | 163.1 | 159.0 | 167.3 | 168.2 | 164.08 | 20.03 | 113.9 | 119.7 | 115.8 | 111.4 | 113.3 | 109.9 | 570.1 | 114.0 |
| червень | 172.5 | 176.9 | 183.0 | 178.9 | 187.1 | 179.68 | 35.63 | 124.7 | 126.8 | 125.6 | 128.2 | 121.2 | 122.3 | 624.1 | 124.8 |
| липень | 168.0 | 171.0 | 177.4 | 177.8 | 182.5 | 175.34 | 31.29 | 121.7 | 123.5 | 121.4 | 124.3 | 120.5 | 119.3 | 609.0 | 121.8 |
| серпень | 159.0 | 162.7 | 163.8 | 168.9 | 173.2 | 165.52 | 21.47 | 114.9 | 116.9 | 115.6 | 114.8 | 114.4 | 113.2 | 574.9 | 115.0 |
| вересень | 135.7 | 139.5 | 144.4 | 147.0 | 153.1 | 143.94 | -0.11 | 99.9 | 99.8 | 99.1 | 101.2 | 99.6 | 100.0 | 499.7 | 99.9 |
| жовтень | 118.9 | 123.2 | 123.3 | 131.8 | 133.9 | 126.22 | -17.83 | 87.6 | 87.4 | 87.5 | 86.4 | 89.3 | 87.5 | 439.1 | 87.6 |
| листопад | 110.0 | 112.6 | 113.8 | 124.8 | 127.4 | 117.72 | -26.33 | 81.7 | 80.9 | 80.0 | 79.7 | 84.6 | 83.3 | 408.5 | 81.7 |
| грудень | 118.8 | 122.5 | 124.8 | 134.8 | 142.0 | 128.58 | -15.47 | 89.3 | 87.4 | 87.0 | 87.5 | 91.3 | 92.8 | 446.0 | 89.2 |
| всього | 1631.9 | 1689.9 | 1712.9 | 1771.6 | 1836.4 | 1728.54 | — | 1200 | 1200 | 1200 | 1200 | 1200 | 1200 | — | 1200 |
| середній
рівень
ряду \bar{y} | 136.0 | 140.8 | 142.7 | 147.6 | 153.0 | 144.05 | — | 100 | 100 | 100 | 100 | 100 | 100 | — | 100 |

- для кожного місяця визначається середня за 5 років (\bar{y}_t):

$$\bar{y}_1 = \frac{117.8 + 125.1 + 126.8 + 131.8 + 138.7}{5} = 128.04,$$

$$\dots$$

$$\bar{y}_{12} = \frac{118.8 + 122.5 + 124.8 + 134.8 + 142.0}{5} = 128.58;$$

- визначається середньомісячний рівень за 5 років:

$$\bar{y} = \frac{128.04 + \dots + 128.58}{12} = \frac{1728.54}{12} = 144.05;$$

- ланки сезонної хвилі дорівнюють $\bar{y}_t - \bar{y}$:

$$128.04 - 144.05 = -16.01,$$

$$\dots$$

$$128.58 - 144.05 = -15.47.$$

2. Метод відношень щомісячних середніх (\bar{y}_t) до середньої за весь період.

$$I_s = \frac{\bar{y}_t}{\bar{y}} \cdot 100 \% \text{ — індекс сезонності,}$$

де \bar{y}_t — середня за кожен місяць, \bar{y} — середньомісячний рівень за весь період.

$$I_{s_1} = \frac{128.04}{144.05} \cdot 100\% = 88.9\%,$$

...

$$I_{s_{12}} = \frac{128.58}{144.05} \cdot 100\% = 89.3\%.$$

3. Метод відношень щомісячних рівнів до середньої місячної даного року.

- для кожного року розраховується середня величина показника за місяць:

$$\bar{y}_{2013} = \frac{117.8 + \dots + 118.8}{12} = 136.0,$$

...

$$\bar{y}_{2017} = \frac{138.7 + \dots + 142.0}{12} = 153.0;$$

- визначається відношення кожного щомісячного фактичного рівня до цих середніх $y_t/\bar{y}_{\text{рік}}$:

$$\frac{117.8}{136.0} \cdot 100\% = 86.6\%,$$

$$\frac{108.1}{136.0} \cdot 100\% = 79.5\%,$$

...

$$\frac{127.4}{153.0} \cdot 100\% = 83.3\%,$$

$$\frac{142.0}{153.0} \cdot 100\% = 92.8\%;$$

- розраховуємо суму по місяцям за 5 років:

$$\text{січень: } 86.6 + 88.8 + 88.8 + 89.3 + 90.7 = 444.2,$$

...

$$\text{грудень: } 87.4 + 87.0 + 87.5 + 91.3 + 92.8 = 446.0.$$

Тоді

$$I_{s \text{ січ.}} = \frac{444.2}{5} = 88.8 \%,$$

...

$$I_{s \text{ груд.}} = \frac{446}{5} = 89.2 \%.$$

При наявності тенденції до збільшення або зменшення рівнів із року в рік застосовуються і інші способи виміру індексів сезонності.

4. Метод відносних величин.

| місяць | $\frac{y_t}{y_{t-1}}$ | $\frac{y_t}{y_{t-1}}$ | $\frac{y_t}{y_{t-1}}$ | $\frac{y_t}{y_{t-1}}$ | $\frac{y_t}{y_{t-1}}$ | \bar{y}_t | скори-
говані
\bar{y}_t | поправка | скориговані
середні
з поправкою | I_s |
|----------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-------------|---------------------------------|----------|---------------------------------------|-------|
| | 2013 | 2014 | 2015 | 2016 | 2017 | | | | | |
| | % | % | % | % | % | | | | | |
| січень | — | 105.3 | 103.6 | 105.7 | 109.9 | 106.1 | 100.0 | — | — | — |
| лютий | 91.8 | 94.6 | 93.2 | 91.9 | 93.6 | 93.0 | 93.0 | 0.54 | 92.5 | 84.5 |
| березень | 118.3 | 115.3 | 117.2 | 117.6 | 115.3 | 116.7 | 108.5 | 1.08 | 107.4 | 98.1 |
| квітень | 103.5 | 101.5 | 100.9 | 101.4 | 100.8 | 101.6 | 110.2 | 1.62 | 108.6 | 99.2 |
| травень | 122.9 | 117.8 | 113.7 | 115.7 | 111.5 | 116.3 | 128.2 | 2.16 | 126.0 | 115.1 |
| червень | 105.9 | 108.5 | 115.1 | 106.9 | 111.3 | 109.5 | 140.4 | 2.70 | 137.7 | 125.7 |
| липень | 97.4 | 96.6 | 96.9 | 99.4 | 97.5 | 97.6 | 137.0 | 3.24 | 133.6 | 122.0 |
| серпень | 94.6 | 95.1 | 92.3 | 95.0 | 94.9 | 94.4 | 129.3 | 3.78 | 125.5 | 114.6 |
| вересень | 85.4 | 85.8 | 88.2 | 86.9 | 88.4 | 86.9 | 112.4 | 4.32 | 108.1 | 97.8 |
| жовтень | 87.6 | 88.3 | 85.3 | 89.7 | 87.4 | 87.7 | 98.6 | 4.86 | 93.7 | 85.6 |
| листопад | 92.5 | 91.4 | 92.3 | 94.7 | 95.1 | 93.2 | 91.9 | 5.40 | 86.5 | 78.9 |
| грудень | 107.9 | 108.9 | 109.6 | 108.0 | 111.5 | 109.2 | 100.4 | 5.94 | 94.5 | 86.3 |
| січень | | | | | | 106.1 | 106.5 | 6.48 | 100.0 | 91.3 |
| всього | | | | | | | | | 1314.1 | 1200 |
| середня | | | | | | | | | 109.5 | 100 |

- визначаємо ланцюгові темпи зростання $\frac{y_t}{y_{t-1}} \cdot 100 \%$:

$$\frac{108.1}{117.8}, \frac{127.9}{108.1}, \frac{132.4}{127.9}, \dots, \frac{118.8}{110.0}, \frac{125.1}{118.8}, \dots;$$

- визначається середня для кожного місяця:

$$\text{січ.} \quad \frac{105.3 + 103.6 + 105.7 + 109.9}{4} = 106.1,$$

$$\text{лют.} \quad \frac{91.8 + 94.6 + 92.3 + 91.4 + 93.6}{5} = 93.0,$$

...

- розрахунок скоригованих середніх (на основі переходу від ланцюгових індексів

до базисних):

$$\begin{aligned}\bar{y}_{\text{січ.}} &= 100\%, \\ \bar{y}_{\text{лют.}} &= 100 \cdot 93.0 : 100 = 93, \\ \bar{y}_{\text{бер.}} &= 93 \cdot 116.7 : 100 = 108.5, \\ \bar{y}_{\text{кв.}} &= 108.5 \cdot 101.6 : 100 = 110.2, \\ &\dots \\ \bar{y}_{\text{січ.}} &= 100.4 \cdot 106.1 : 100 = 106.5;\end{aligned}$$

- різниця (дивимось січень місяць): $106.5 - 100 = 6.5$.

Поправка: $\frac{6.5}{12} \cdot (i - 1) = 0.54$, i — номер місяця в році, для січня $i = 12$, так як за січень відбулось збільшення, то $\frac{6.5}{12} = 0.54$ беремо зі знаком мінус (-).
Поправка для січня: $(-0.54) \cdot 12 \approx -6.5$.

Скориговані середні для кожного місяця з урахуванням поправок:

$$\begin{aligned}\text{лют.} \quad &93.0 - 0.54 \approx 92.5, \\ \text{бер.} \quad &108.5 - 2 \cdot 0.54 \approx 107.4, \\ &\dots \\ \text{січ.} \quad &106.5 - 6.5 = 100;\end{aligned}$$

- далі розраховуємо середню скоригованих середніх для кожного місяця з урахуванням поправок:

$$\frac{92.5 + 107.4 + \dots + 100}{12} = 109.5;$$

- співставимо скориговані середні з середньою 109.5:

$$\text{лют. : } 92.5 : 109.5 \cdot 100\% = 84.5\%, \quad \dots \quad \text{січ. : } 100 : 109.5 \cdot 100\% = 91.3\%.$$

5. Метод відносних величин на основі медіани.

| місяць | $\frac{y_t}{y_{t-1}}$ | $\frac{y_t}{y_{t-1}}$ | $\frac{y_t}{y_{t-1}}$ | $\frac{y_t}{y_{t-1}}$ | $\frac{y_t}{y_{t-1}}$ | ранжировані темпи зростання | | | | | M, % | скори-
говані
M | поп-
равка
Δ | скориг.
ранжир.
з урахув.
попр. | I _s |
|----------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------------|-------|-------|-------|-------|--------|-----------------------|--------------------|--|----------------|
| | 2013 | 2014 | 2015 | 2016 | 2017 | I | II | III | IV | V | | | | | |
| | % | % | % | % | % | | | | | | | | | | |
| січень | — | 105.3 | 103.6 | 105.7 | 109.9 | — | 103.6 | 105.3 | 105.7 | 109.9 | 105.5 | 100 | — | — | — |
| лютий | 91.8 | 94.6 | 93.2 | 91.9 | 93.6 | 91.8 | 91.9 | 93.2 | 93.6 | 94.6 | 93.2 | 93.2 | 0.325 | 92.9 | 84.4 |
| березень | 118.3 | 115.3 | 117.2 | 117.6 | 115.3 | 115.3 | 115.7 | 117.2 | 117.6 | 118.3 | 117.2 | 109.2 | 0.650 | 100.6 | 98.6 |
| квітень | 103.5 | 101.5 | 100.9 | 101.4 | 100.8 | 100.8 | 100.9 | 101.4 | 101.5 | 103.5 | 101.4 | 110.7 | 0.975 | 105.7 | 99.6 |
| травень | 122.9 | 117.8 | 113.7 | 115.7 | 111.5 | 111.5 | 113.7 | 115.7 | 117.8 | 122.9 | 115.7 | 128.1 | 1.300 | 126.8 | 115.2 |
| червень | 105.9 | 108.5 | 115.1 | 106.9 | 111.3 | 105.9 | 106.9 | 108.5 | 111.3 | 115.1 | 108.5 | 139.0 | 1.625 | 137.4 | 124.8 |
| липень | 97.4 | 96.6 | 96.9 | 99.4 | 97.5 | 96.6 | 96.9 | 97.4 | 97.5 | 99.4 | 97.4 | 135.4 | 1.950 | 133.4 | 121.2 |
| серпень | 94.6 | 95.1 | 92.3 | 95.0 | 94.9 | 92.3 | 94.6 | 94.9 | 95.0 | 95.1 | 94.9 | 128.5 | 2.275 | 126.2 | 114.6 |
| вересень | 85.4 | 85.8 | 88.2 | 86.9 | 88.4 | 85.4 | 85.8 | 86.9 | 88.2 | 88.4 | 86.9 | 111.7 | 2.600 | 109.1 | 99.1 |
| жовтень | 87.6 | 88.3 | 85.3 | 89.7 | 87.4 | 85.3 | 87.4 | 87.6 | 88.3 | 89.7 | 87.6 | 97.8 | 2.925 | 94.9 | 86.2 |
| листопад | 92.5 | 91.4 | 92.3 | 94.7 | 95.1 | 91.4 | 92.3 | 92.5 | 94.7 | 95.1 | 92.5 | 90.5 | 3.290 | 87.2 | 79.2 |
| грудень | 107.9 | 108.9 | 109.6 | 108.0 | 111.5 | 107.9 | 108.0 | 108.9 | 109.6 | 111.5 | 108.9 | 98.5 | 3.575 | 94.9 | 86.2 |
| січень | | | | | | | | | | | 105.5 | 103.9 | 3.900 | 100 | 90.8 |
| всього | | | | | | | | | | | 1208.6 | 1321.1 | | | |
| середня | | | | | | | | | | | 110.1 | | | | |

- Визначаються ланцюгові темпи зростання щомісячно.
- Ланцюгові $T_{i(\text{зр})}$ ранжуються у порядку зростання.
- Визначається медіана по щомісячним даним:

$$M_{\text{січ.}} = \frac{105.3 + 105.7}{2} = 105.5, \quad M_{\text{лют.}} = 93.2, \quad \dots$$

- Коригуються медіани:

$$M_{\text{січ.}} = 100\%, \quad M_{\text{лют.}} = 93.2 \cdot 100 : 100 = 93.2, \quad M_{\text{бер.}} = 93.2 \cdot 117.2 : 100 = 109.2, \\ M_{\text{кв.}} = 109.2 \cdot 101.4 : 100 = 110.7, \quad \dots \quad M_{\text{січ.}} = 98.5 \cdot 105.5 : 100 = 103.9.$$

- Поправка: $\Delta = (103.9 - 100) : 12 = 0.325$, так як $M_{\text{січ.}}$ збільшилось, то поправку віднімаємо.
- Скориговані M з урахуванням поправки:

$$M_{\text{лют.}} = 93.2 - 0.325 = 92.9, \quad M_{\text{бер.}} = 109.2 - 2 \cdot 0.325 = 108.6, \\ M_{\text{кв.}} = 110.7 - 3 \cdot 0.325 = 109.7, \quad \dots \quad M_{\text{січ.}} = 103.9 - 0.325 \cdot 12 = 100.$$

- Розраховуємо середню скоригованих поправкою M :

$$(92.9 + 108.6 + 109.7 + \dots + 100) : 12 = 110.1.$$

- Співставляємо скориговані поправкою значення M із середньою:

$$\frac{92.9}{110.1} = 84.4, \quad \frac{108.6}{110.1} = 98.6, \quad \dots \quad \frac{100}{110.1} = 90.8.$$

Отримали $I_s = -; 84.4; 98.6; 99.6; 115.2; 124.8; 121.2; 114.6; 99.1; 86.2; 79.2; 86.2; 90.8$.

3.1 Побудова аналітичного вигляду сезонної компоненти

У багатьох випадках моделювання рядів динаміки за допомогою поліномів або інших функцій не дає бажаних результатів, тому що у ряді динаміки є помітне періодичне коливання навколо загальної тенденції або спостерігається автокореляція не у самих рівнях, а у їх відхиленнях від отриманих за певними аналітичними формулами теоретичних значень. У таких випадках використовується гармонічний аналіз.

Апроксимація динаміки економічних явищ рядом Фур'є полягає у виборі таких гармонічних коливань, накладання яких один на одного відображало б періодичні коливання фактичних рівнів часового ряду. За допомогою ряду Фур'є можна представити динаміку явищ у вигляді деякої функції часу t , в якій доданки розташовані за зменшенням періодів.

Якщо кількість рівновіддалених точок парна і дорівнює $2N$, то розклад функції $f(x)$ у ряд Фур'є має вигляд

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^M \left(a_k \cos \frac{\pi}{N} kx + b_k \sin \frac{\pi}{N} kx \right) + \frac{a_N}{2} \cos \pi x, \quad M \leq N - 1,$$

де

$$a_k = \frac{1}{N} \sum_{l=0}^{2N-1} f\left(\frac{\pi}{N} l\right) \cos \frac{\pi}{N} kl, \quad k = \overline{0, N},$$

$$b_k = \frac{1}{N} \sum_{l=0}^{2N-1} f\left(\frac{\pi}{N} l\right) \sin \frac{\pi}{N} kl, \quad k = \overline{1, N-1}.$$

Параметр M визначає кількість гармонік.

Для непарної $(2N + 1)$ кількості рівновіддалених точок ряд Фур'є має вигляд

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^M \left(a_k \cos \frac{2\pi}{2N+1} kx + b_k \sin \frac{2\pi}{2N+1} kx \right), \quad M \leq N - 1,$$

де

$$a_k = \frac{2}{2N+1} \sum_{l=0}^{2N} f\left(\frac{2\pi}{2N+1} l\right) \cos \frac{2\pi}{2N+1} kl, \quad k = \overline{0, N},$$

$$b_k = \frac{2}{2N+1} \sum_{l=0}^{2N} f\left(\frac{2\pi}{2N+1} l\right) \sin \frac{2\pi}{2N+1} kl, \quad k = \overline{1, N}.$$

Оцінка адекватності і точності трендових моделей

4.1 Оцінка адекватності моделі

Незалежно від того, як ми будували модель, питання про можливість застосування її для аналізу і прогнозування економічних показників можливе тільки після встановлення її адекватності.

Адекватність моделі еквівалентна таким вимогам до випадкової залишкової величини:

- 1) випадковість коливань рівнів залишкової величини;
- 2) відповідність розподілу випадкової залишкової величини нормальному закону розподілу;
- 3) рівність математичного очікування випадкової залишкової величини нулю;
- 4) незалежність значень рівнів випадкової залишкової величини;

Розглянемо кожну вимогу окремо.

4.1.1 Перевірка випадковості коливань рівнів залишкової величини

Для перевірки випадковості $\varepsilon_t = y_t - \hat{y}_t$ ($t = \overline{1, n}$) можна скористатися розглянутими раніше тестами: критерій серій, заснований на медіані; критерій піків і критерій Аббе.

4.1.2 Перевірка відповідності розподілу випадкової компоненти нормальному розподілу

Критерій Девіда — Хартлі — Пірсона (RS-критерій)

Критерій нормальності розподілу ймовірності випадкової величини ґрунтується на розподілі відношення розкиду до стандартного відхилення.

Статистика критерію має вигляд

$$U = \frac{R}{S},$$

де $R = y_{\max} - y_{\min}$, S — стандартне відхилення.

Гіпотеза нормальності приймається, якщо $U_1(\alpha) < U < U_2(\alpha)$ (α — рівень значущості).

Критерій нормальності Фроціні

Фроціні запропонував простий, але достатньо потужний критерій нормальності з параметрами, що оцінюються за вибіркою, і ґрунтується на статистиці

$$B_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left| \Phi(z_i) - \frac{i - 0.5}{n} \right|,$$

де $z_i = \frac{y_i - \bar{y}}{s}$; $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$; $s^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$; $\Phi(z_i)$ — функція розподілу $N(0, 1)$.

Критичні значення статистики B_n наведені в [табл. В.15](#).

Якщо $B_n < B_n(\alpha)$, то гіпотеза про нормальність розподілу випадкових величин не відхиляється.

4.1.3 Перевірка рівності математичного очікування випадкової компоненти нулю

Якщо випадкова величина розподілена за нормальним законом, то перевірку на рівність математичного очікування випадкової величини нулевій здійснимо за допомогою t -критерію Ст'юдента. Розрахункове значення цього критерію розраховується за формулою

$$t_{\text{СТ}} = \frac{\bar{\varepsilon} - 0}{S_{\varepsilon}} \sqrt{n},$$

де $\bar{\varepsilon}$ — середнє арифметичне значення рівнів залишкової послідовності ε_t , S_{ε} — стандартне відхилення для цієї послідовності.

Якщо $|t_{CT}| < t_{KP} = t(\alpha; n - 1)$, то гіпотеза про рівність нулю математичного очікування випадкової послідовності приймається, в іншому випадку модель вважається неадекватною.

4.1.4 Перевірка незалежності значень рівнів випадкової компоненти

Якщо між залишками існує авторегресійний процес першого порядку, тобто

$$\varepsilon_t = \rho \varepsilon_{t-1} + u_t,$$

де $|\rho| < 1$, $M(u_t) = 0$ для всіх t , $\sigma_{u_t u_{t''}} = \begin{cases} 0, & t' \neq t'', \\ \sigma_u^2, & t' = t'', \end{cases}$ то його можна виявити за допомогою тесту Дарбіна — Уотсона.

Тест Дарбіна — Уотсона

1) Формулюються гіпотези:

$H_A : \rho \neq 0$, присутня автокореляція,

$H_0 : \rho = 0$, відсутня автокореляція.

2) Задаємо рівень значущості α .

3) Розраховуємо d_{CT} за формулою

$$d_{CT} = \frac{\sum_{i=2}^n (\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=1}^n \varepsilon_i^2}.$$

Параметр d_{CT} належить проміжку $[0; 4]$. Так як тест двосторонній, d_{KP} знаходимо за таблицями при $\alpha := \frac{\alpha}{2}$.

4) Нижче представлено області прийняття рішень при d -тесті нульової гіпотези $H_0 : \rho = 0$.

| | | | | | | |
|---|-------|-------|---|-----------|-----------|---|
| 0 | d_0 | d_u | 2 | $4 - d_u$ | $4 - d_0$ | 4 |
| – | ? | + | ? | – | | |

Умовні позначення:

| | |
|---|-----------------------------------|
| – | області відхилення гіпотези H_0 |
| ? | області невизначеності |
| + | область прийняття гіпотези H_0 |

Якщо d_{CT} попадає в область невизначеності, то приймається гіпотеза про наявність автокореляції, хоча вона може бути і відсутня. Так як тест Дарбіна — Уотсона може виявити автокореляцію тільки першого порядку, а вона може бути і вищих порядків, то краще користуватись тестом серій Бреуша — Годфрі.

Тест Бреуша — Годфрі

Нехай існує автокореляція залишків p -го порядку, тобто

$$\varepsilon_t = \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2} + \dots + \alpha_p \varepsilon_{t-p} + u_t, \quad (*)$$

де u_t — залишок у виписаному регресійному рівнянні. Методом найменших квадратів оцінюємо коефіцієнти α_i . Якщо модель (*) значимо відрізняється від нуля, то існує автокореляція. Для визначення значущості моделі розрахуємо

$$F_{CT} = \frac{\widehat{\varepsilon}^\top \widehat{\varepsilon} (n-p)}{\widehat{u}^\top \widehat{u} (p-1)},$$

де $\widehat{\varepsilon}$ — оцінені значення залишків, $\widehat{u}_i = \varepsilon_i - \widehat{\varepsilon}_i$. Якщо $F_{CT} < F_{KP} = F(\alpha; p-1; n-p)$, то модель (*) значимо відрізняється від нуля. На практиці найчастіше $p = 1, 2, 3$ або 4 .

Отже, якщо всі чотири властивості залишкової послідовності виконані, то модель адекватна.

4.2 Оцінка точності моделі

Якщо модель адекватна, то є сенс розглядати задачу оцінки її точності. У якості статистичних показників точності використовуються такі:

1) Середнє квадратичне відхилення

$$\sigma_{\varepsilon} = \sqrt{\frac{1}{n-p} \sum_{t=1}^n (y_t - \hat{y}_t)^2}.$$

За цим параметром порівнюються між собою декілька моделей, і та буде кращою, у якої середнє квадратичне відхилення менше.

2) Середня відносна похибка апроксимації

$$MAPE = \frac{100\%}{n} \sum \frac{|y_t - \hat{y}_t|}{y_t}.$$

Ця величина має такі порогові значення:

| Оцінка MAPE, % | Характеристика якості моделі |
|----------------|------------------------------------|
| < 10 | висока точність апроксимації |
| 10 ÷ 20 | добра точність апроксимації |
| 20 ÷ 50 | задовільна точність апроксимації |
| > 50 | незадовільна точність апроксимації |

3) Коефіцієнт детермінації

$$R^2 = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{\hat{y}})^2}{\sum_{t=1}^n (y_t - \bar{y})^2},$$

де $\bar{\hat{y}}$ — середнє арифметичне оцінених значень y за моделлю, \bar{y} — середнє арифметичне часового ряду. Коефіцієнт детермінації змінюється від 0 до 1, тобто $0 \leq R^2 \leq 1$. Чим ближче R^2 до 1, тим точніша модель.

Коефіцієнт детермінації має недолік: якщо в моделі збільшується кількість параметрів, то R^2 збільшується. Тому на практиці досить часто користуються скоригованим коефіцієнтом детермінації за Тейлом:

$$\bar{R}_T^2 = 1 - (1 - R^2) \frac{n-1}{n-p},$$

де p — кількість параметрів у моделі.

Приклад 4.1.

| t | факт.
знач.
y_t | розрах.
\hat{y}_t | відхи-
лення
ε_t | | ε_t^2 | $\varepsilon_t - \varepsilon_{t-1}$ | $(\varepsilon_t - \varepsilon_{t-1})^2$ | $\frac{ \varepsilon_t }{y_t} 100\%$ |
|----------|-------------------------|------------------------|------------------------------------|---|-------------------|-------------------------------------|---|-------------------------------------|
| 1 | 85 | 84.4 | 0.6 | | 0.36 | — | — | 0.71 |
| 2 | 81 | 81.0 | 0.0 | — | 0.00 | —0.6 | 0.36 | 0.00 |
| 3 | 78 | 77.6 | 0.4 | + | 0.16 | 0.4 | 0.16 | 0.49 |
| 4 | 72 | 74.1 | -2.1 | — | 4.41 | -2.5 | 6.25 | 2.69 |
| 5 | 69 | 70.7 | -1.7 | + | 2.89 | 0.4 | 0.16 | 2.46 |
| 6 | 70 | 67.3 | 2.7 | + | 7.29 | 4.4 | 19.36 | 3.86 |
| 7 | 64 | 63.8 | 0.2 | — | 0.04 | -2.5 | 6.25 | 0.31 |
| 8 | 61 | 60.4 | 0.6 | + | 0.36 | 0.4 | 0.16 | 0.98 |
| 9 | 56 | 57.0 | -1.0 | — | 1.00 | -1.6 | 2.56 | 1.79 |
| Σ | 636 | 636.3 | -0.3 | + | 15.51 | | 35.26 | 13.29 |

Для часового ряду (другий стовпчик) побудована трендова модель у вигляді многочлена першого ступеня:

$$\hat{y}_t = 87.8 - 3.4t.$$

Необхідно оцінити адекватність і точність побудованої моделі.

Перевірку випадковості рівнів ряду зробимо на основі тесту «висхідних» і «спадних» серій. Маємо $n = 9$, $\nu(9) = 7$, $\left[\frac{1}{3}(2n-1) - 1.96\sqrt{\frac{16n-29}{90}}\right] = 3$, $\tau(9) = 2$, $\tau_0(9) = 5$, $\nu(9) = 7 > 3$, $\tau(9) = 2 < \tau_0(9) = 5$.

Так як нерівності виконуються, то ряд відхилень випадковий.

Для перевірки ряду ε_t на нормальність скористаємося критерієм Девіда — Хар-тлі — Пірсона. Маємо: $R = \varepsilon_{\max} - \varepsilon_{\min} = 2.7 - (-2.1) = 4.8$, $S_\varepsilon = \sqrt{\frac{\sum_{t=1}^n \varepsilon_t^2}{n-1}} = \sqrt{\frac{15.51}{8}} = 1.39$. Тоді $U = \frac{R}{S} = 3.45$. Значення U попадає в інтервал $U_1 < U < U_2$ при $\alpha = 0.05$ і $n = 9$: $2.590 < 3.45 < 3.552$. Отримали, що ряд відхилень підпорядкований нормальному закону розподілу.

Перейдемо до перевірки рівності нулю математичного очікування ряду відхилень ε_t . Для цього розрахуємо

$$t_{\text{СТ}} = \frac{\bar{\varepsilon}}{S_\varepsilon} \sqrt{n} = \frac{0.033}{1.39} \sqrt{9} = 0.07, \quad t_{\text{КР}} = t(0.05; 9-1) = t(0.05; 8) = 2.306.$$

Так як $t_{\text{СТ}} < t_{\text{КР}}$, то гіпотеза про рівність нулю математичного сподівання випадкової послідовності приймається з ймовірністю 95%.

Для перевірки незалежності значень рівнів випадкової компоненти скористає-

мося тестом Дарбіна — Уотсона. Розрахуємо $d_{CT} = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=2}^n \varepsilon_t^2} = \frac{35.26}{15.51} = 2.27$. Маємо при $\alpha := 0.1$

| | | | | | | |
|---|------|------|---|------|------|---|
| 0 | 0.82 | 1.32 | 2 | 2.68 | 3.18 | 4 |
| — | ? | + | ? | — | | |

d_{CT} попадає в область прийняття гіпотези про відсутність автокореляції першого порядку відхилень ε_t .

З отриманого виходить, що всі чотири умови адекватності моделі виконуються.

Для характеристики точності моделі розрахуємо параметр $MAPE$ — середню відносну похибку апроксимації:

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \frac{\varepsilon_t}{y_t} = \frac{13.29}{9} \% = 1.48 \%$$

Значення $MAPE = 1.48\%$ свідчить про відмінну апроксимацію.

Залишилось розрахувати коефіцієнт детермінації R^2 . Маємо $R^2 = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2} = 0.9659$. Якщо розрахувати скоригований коефіцієнт детермінації за Тейлом, то отримаємо $\bar{R}_T^2 = 1 - (1 - R^2) \frac{n-1}{n-p} = 1 - (1 - 0.97) \frac{9-1}{9-2} = 0.9610$. Коефіцієнт детермінації досить близький до 1, а це означає, що модель адекватна. ☒

Прогнозування часових рядів

5.1 Найпростіші методи прогнозування

Найбільш простими методами прогнозування є такі:

- 1) прогнозування в припущенні, що присутня абсолютна незмінність значень попередніх рівнів у майбутньому;
- 2) метод середнього рівня ряду;
- 3) метод середнього абсолютного приросту;
- 4) метод середнього темпу зросту.

5.1.1 Прогнозування у припущенні незмінності значень попередніх рівнів

Це означає, що стверджується, що кожне наступне прогнозне значення буде дорівнювати попередньому значенню рівня, тобто: $\hat{y}_{t+L} = \hat{y}_{t+L-1}$.

Такий випадок дуже рідко зустрічається у практиці прогнозування. Його ще називають «наївним».

5.1.2 Прогнозування на підставі середнього рівня ряду

Якщо середній рівень ряду не має тенденції до зміни і якщо зміни незначні (ряд стаціонарний), то приймаємо

$$y_{t+L} = \bar{y}.$$

Це означає, що прогнозний рівень дорівнює середньому значенню рівнів у минулому. Отриманий прогноз є точковим. Але ряд можна розглядати як вибірку з якоїсь генеральної сукупності і тому середня вибірки буде змінюватися і доцільно розрахувати прогнозний інтервал. Він має такі межі:

$$\text{Н. М.: } \hat{y}_{t+L} - t(\alpha; n-1)\sigma_y \sqrt{1 + \frac{1}{n}},$$

$$\text{В. М.: } \hat{y}_{t+L} + t(\alpha; n-1)\sigma_y \sqrt{1 + \frac{1}{n}},$$

$$\text{де } \sigma_y = \sqrt{\frac{\sum_t^n (y_t - \bar{y})^2}{n-1}}.$$

Середню краще брати не по всьому ряду, а по правій його частині у випадку, якщо спостерігається хоча б і мала тенденція.

5.1.3 Прогнозування методом середнього абсолютного приросту

Запропонований метод використовується у випадку лінійної тенденції і виконання таких передумов:

- 1) абсолютні ланцюгові прирости ($\Delta_t = y_t - y_{t-1}$) приблизно однакові;
- 2) виконується нерівність $\sigma_{\text{зал.}}^2 \leq \rho^2$, де $\sigma_{\text{зал.}}^2$ — залишкова дисперсія визначається за формулою $\sigma_{\text{зал.}}^2 = \frac{\sum_{t=1}^n (y_t - \bar{y}_\Delta)^2}{n}$, де y_t — емпіричні значення рівнів ряду, \bar{y}_Δ — теоретичні значення рівнів ряду, вирівнених методом середнього абсолютного приросту, ρ^2 визначається так:

$$\rho^2 = \frac{1}{2} \frac{\sum \Delta_t^2}{n}.$$

Загальна модель прогнозу має вигляд:

$$\hat{y}_{t+L} = y_t + \bar{\Delta} \times L,$$

де y_t — останній рівень ряду, L — період випередження прогнозу, $\bar{\Delta} = \frac{y_n - y_1}{n - 1}$ — середній абсолютний приріст.

Приклад 5.1. За даними про об'єм споживання електроенергії у місті за січень-серпень побудувати прогноз методом середнього абсолютного приросту на вересень і жовтень.

| місяць | кВт/год | Абсолютні | | | | |
|----------|---------|--------------------|------------------|------------------------|----------------------------|--------------|
| | | ланцюгові прирости | \bar{y}_Δ | $y_t - \bar{y}_\Delta$ | $(y_t - \bar{y}_\Delta)^2$ | Δt^2 |
| січень | 201.8 | — | 201.8 | 0 | 0 | — |
| лютий | 202.4 | 0.6 | 202.8 | −0.4 | 0.16 | 0.36 |
| березень | 203.1 | 0.7 | 203.8 | −0.7 | 0.49 | 0.49 |
| квітень | 204.0 | 0.9 | 204.8 | −0.8 | 0.64 | 0.81 |
| травень | 205.2 | 1.2 | 205.8 | −0.6 | 0.36 | 1.44 |
| червень | 206.4 | 1.2 | 206.8 | −0.4 | 0.16 | 1.44 |
| липень | 207.6 | 1.2 | 207.8 | −0.2 | 0.04 | 1.44 |
| серпень | 208.8 | 1.2 | 208.8 | 0 | 0 | 1.44 |
| Всього | 1639.3 | 7.0 | — | — | 1.85 | 7.42 |

$$\bar{\Delta} = \frac{y_n - y_1}{n - 1} = \frac{208.8 - 201.8}{8 - 1} = 1.$$

Далі перевіримо нерівність $\sigma_{\text{зал.}}^2 \leq \rho^2$. Маємо

$$\sigma_{\text{зал.}}^2 = \frac{\sum_{t=1}^n (y_t - \bar{y}_\Delta)^2}{n} = \frac{1.85}{8} = 0.231, \quad \rho^2 = \frac{1}{2} \sum_{t=1}^n \Delta t^2 = \frac{1}{2} \cdot \frac{7.42}{8} = 0.464.$$

Отримали нерівність $\sigma_{\text{зал.}}^2 < \rho^2$, і тому можна робити прогноз:

$$\hat{y}_{\text{вер.}} = 208.8 + 1 \cdot 1 = 209.8 \text{ кВт/год},$$

$$\hat{y}_{\text{жов.}} = 208.8 + 1 \cdot 2 = 210.8 \text{ кВт/год}.$$



5.1.4 Прогнозування методом середнього темпу зросту

Цим методом роблять прогноз у випадку, якщо тенденція характеризується експоненціальною кривою. Прогноз методом середнього темпу зросту робиться за формулою

$$\hat{y}_{t+L} = y_t \times \bar{T}_p,$$

$$\bar{T}_p = \sqrt[n-1]{\frac{y_n}{y_1}} - \text{середній темп росту.}$$

Зазначимо, що сума емпіричних значень початкового ряду повинна дорівнювати сумі значень ряду, вирівнених за середнім темпом росту. Якщо ці суми не співпадають, то це означає:

- 1) початковий ряд має не експоненціальну тенденцію;
- 2) випадкові фактори мають великий вплив на формування ряду.

Приклад 5.2. За даними споживання електроенергії у місті розрахувати прогноз на червень і липень методом середнього темпу зросту.

| млн кВт/год | T_p — | |
|-------------|------------|------------------------|
| | ланцюговий | $\hat{y}_{\text{тр.}}$ |
| | темپ росту | |
| 10 | — | 10.0 |
| 11 | 1.10 | 11.4 |
| 13 | 1.18 | 13.0 |
| 15 | 1.15 | 14.8 |
| 17 | 1.13 | 16.8 |
| 66 | — | 66 |

Середній темп росту такий:

$$\bar{T}_p = \sqrt[n-1]{\frac{y_n}{y_1}} = \sqrt[5-1]{\frac{17}{10}} = 1.14.$$

Тоді прогноз споживання електроенергії у місті такий:

$$\begin{aligned}\hat{y}_{\text{червень}} &= 17 \cdot 1.14^1 = 19.4 \text{ млн кВт/год,} \\ \hat{y}_{\text{липень}} &= 17 \cdot 1.14^2 = 22.1 \text{ млн кВт/год.}\end{aligned}$$



Розглянуті методи прогнозування є найпростішими і прогнози, зроблені за ними, не дуже точні. Тому ці методи використовуються тільки для короткострокових прогнозів.

5.2 Прогнозування динаміки за трендовими моделями

Прогнозування економічних показників за трендовими моделями засновано на ідеї екстраполяції. Тобто вважається, що дія зв'язків і закономірностей, діючих в сучасному періоді, розповсюджується і на наступні періоди. Застосування екстраполяції доцільне при виконанні таких умов:

- початковий часовий ряд повинен бути довгим;
- часовий ряд не повинен мати стрибків і тенденція ряду описується плавною кривою;
- екстраполяція за допомогою кривих тренду дасть прийнятні результати, якщо границя насичення буде визначена досить точно.

Прогноз на основі трендових моделей включає в собі два моменти: визначення точкового і інтервального прогнозів.

Точковий прогноз — це прогноз, що отримується підкладанням у рівняння тенденції величини часу t , що відповідає відповідному періоду випередження: $t = n + 1, n + 2, \dots$. Зрозуміло, що співпадіння фактичних даних в майбутньому і прогнозних точкових малоімовірне. Тому доводиться розраховувати нижню і верхню межі зміни прогнозної величини. Цей інтервал називають прогнозним інтервалом. Більш точно: прогнозний інтервал — це інтервал з випадковими межами, в якому з ймовірністю $(1 - \alpha) 100\%$ можна очікувати фактичне значення прогнозного показника.

Розрахунок прогнозних інтервалів при прогнозуванні за допомогою трендових моделей спирається на формули теорії регресій. Але слід пам'ятати, що динамічні ряди відрізняються від статистичних сукупностей, і тому необхідно обережно підходити до оцінки прогнозних інтервалів.

Далі розглянемо різні види трендових моделей, оцінки їх параметрів і вкажемо на алгоритми розрахунків прогнозних інтервалів.

5.2.1 Лінійний тренд

У цьому випадку функція тенденції має вигляд

$$y_t = a_0 + a_1 t.$$

Розрахунок параметрів a_0 і a_1 здійснюється методом найменших квадратів, який дає таку систему рівнянь для розрахунку цих параметрів:

$$\begin{aligned} a_0 n + a_1 \sum_{t=1}^n t &= \sum_{t=1}^n y_t, \\ a_0 \sum_{t=1}^n t + a_1 \sum_{t=1}^n t^2 &= \sum_{t=1}^n t \cdot y_t. \end{aligned}$$

У випадку лінійного тренду межі прогнозного інтервалу мають вигляд:

$$\begin{aligned} \text{Н. М.: } \hat{y}_{n+L} - t(\alpha, n-2) \cdot S_{\hat{y}} \sqrt{1 + \frac{1}{n} + \frac{1}{\sum t^2} t_L^2}, \\ \text{В. М.: } \hat{y}_{n+L} + t(\alpha, n-2) \cdot S_{\hat{y}} \sqrt{1 + \frac{1}{n} + \frac{1}{\sum t^2} t_L^2}, \end{aligned}$$

де L — період випередження, \hat{y}_{n+L} — точковий прогноз за трендовою моделлю в момент $(n+L)$, n — кількість спостережень в часовому ряді, $S_{\hat{y}} = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{n-2}}$, $t(\alpha, n-2)$ — табличне значення критерію Ст'юдента для рівня значущості α і кількості ступенів свободи $(n-2)$, \hat{y}_t — розраховані за лінійною функцією.

Для різних значень n , α і L є таблиці, в яких протабульовано вираз $K^* = t(\alpha, n-2) \sqrt{1 + \frac{1}{n} + \frac{1}{\sum t^2} t_L^2}$. Тоді межі прогнозного інтервалу будуть такі:

$$\begin{aligned} \text{Н. М.: } \hat{y}_{n+L} - S_{\hat{y}} \cdot K^*, \\ \text{В. М.: } \hat{y}_{n+L} + S_{\hat{y}} \cdot K^*. \end{aligned}$$

5.2.2 Многочлен другого ступеня

Функція тенденції має вигляд:

$$y_t = a_0 + a_1 t + a_2 t^2.$$

Система рівнянь для визначення параметрів a_0, a_1 і a_2 має вигляд:

$$\begin{aligned} a_0 n + a_1 \sum_{t=1}^n t + a_2 \sum_{t=1}^n t^2 &= \sum_{t=1}^n y_t, \\ a_0 \sum_{t=1}^n t + a_1 \sum_{t=1}^n t^2 + a_2 \sum_{t=1}^n t^3 &= \sum_{t=1}^n t \cdot y_t, \\ a_0 \sum_{t=1}^n t^2 + a_1 \sum_{t=1}^n t^3 + a_2 \sum_{t=1}^n t^4 &= \sum_{t=1}^n t^2 y_t. \end{aligned}$$

Прогнозний інтервал має такі межі:

$$\begin{aligned} \text{Н. М.: } \hat{y}_{n+L} - t(\alpha; n-3) S_{\hat{y}} \sqrt{1 + \frac{t_L^2}{\sum_{t=1}^n t^2} + \frac{\sum_{t=1}^n t^4 - 2t_L^2 \sum_{t=1}^n t^2 + n t_L^4}{n \sum_{t=1}^n t^4 - (\sum_{t=1}^n t^2)^2}}, \\ \text{В. М.: } \hat{y}_{n+L} + t(\alpha; n-3) S_{\hat{y}} \sqrt{1 + \frac{t_L^2}{\sum_{t=1}^n t^2} + \frac{\sum_{t=1}^n t^4 - 2t_L^2 \sum_{t=1}^n t^2 + n t_L^4}{n \sum_{t=1}^n t^4 - (\sum_{t=1}^n t^2)^2}}, \end{aligned}$$

де $S_{\hat{y}}$ розраховується так як і для лінійного тренду, тільки у знаменнику кількість вільностей — $(n-3)$.

Позначимо вираз

$$t(\alpha; n-3) \sqrt{1 + \frac{t_L^2}{\sum_{t=1}^n t^2} + \frac{\sum_{t=1}^n t^4 - 2t_L^2 \sum_{t=1}^n t^2 + n t_L^4}{n \sum_{t=1}^n t^4 - (\sum_{t=1}^n t^2)^2}}$$

через K^* . Існують таблиці табульованих значень K^* для різних n, α і L . Тоді прогнозний інтервал має вигляд:

$$\begin{aligned} \text{Н. М.: } \hat{y}_{n+L} - S_{\hat{y}} \cdot K^*, \\ \text{В. М.: } \hat{y}_{n+L} + S_{\hat{y}} \cdot K^*. \end{aligned}$$

5.2.3 Многочлен третьего степени

Функція тенденції має вигляд:

$$y_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3.$$

Для визначення параметрів a_0 , a_1 , a_2 і a_3 маємо таку систему рівнянь:

$$\begin{aligned} a_0 n + a_1 \sum_{t=1}^n t + a_2 \sum_{t=1}^n t^2 + a_3 \sum_{t=1}^n t^3 &= \sum_{t=1}^n y_t, \\ a_0 \sum_{t=1}^n t + a_1 \sum_{t=1}^n t^2 + a_2 \sum_{t=1}^n t^3 + a_3 \sum_{t=1}^n t^4 &= \sum_{t=1}^n t y_t, \\ a_0 \sum_{t=1}^n t^2 + a_1 \sum_{t=1}^n t^3 + a_2 \sum_{t=1}^n t^4 + a_3 \sum_{t=1}^n t^5 &= \sum_{t=1}^n t^2 y_t, \\ a_0 \sum_{t=1}^n t^3 + a_1 \sum_{t=1}^n t^4 + a_2 \sum_{t=1}^n t^5 + a_3 \sum_{t=1}^n t^6 &= \sum_{t=1}^n t^3 y_t. \end{aligned}$$

Межі прогнозного інтервалу такі:

$$\begin{aligned} \text{Н. М.: } \hat{y}_{n+L} - t(\alpha; n-4) S_{\hat{y}} \sqrt{1 + \frac{1}{\sum_{t=1}^n t^2} t_L^2 + \frac{\sum t^4 - 2t_L^2 \sum t^2 + n t_L^4}{n \sum t^4 - (\sum t^2)^2} + \frac{(\sum t^6 - 2 \sum t^4) t_L^2 + (\sum t^2) t_L^6}{\sum t^2 \sum t^6 - (\sum t^4)^2}}, \\ \text{В. М.: } \hat{y}_{n+L} + t(\alpha; n-4) S_{\hat{y}} \sqrt{1 + \frac{1}{\sum_{t=1}^n t^2} t_L^2 + \frac{\sum t^4 - 2t_L^2 \sum t^2 + n t_L^4}{n \sum t^4 - (\sum t^2)^2} + \frac{(\sum t^6 - 2 \sum t^4) t_L^2 + (\sum t^2) t_L^6}{\sum t^2 \sum t^6 - (\sum t^4)^2}}, \end{aligned}$$

де $S_{\hat{y}} = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n-4}}$, \hat{y}_t — розраховані за многочленом третього ступеня.

5.2.4 Проста експонента

Розглянемо тенденцію виду: $y_t = ab^t$. Логарифмуємо це співвідношення:

$$\ln y_t = \ln a + t \ln b.$$

Отримали лінійну функцію відносно t . Параметри a і b розраховуємо так, як і у випадку лінійного тренду. Для спрощеного розрахунку інтервалу можна скористатися параметром K^* для лінійного тренду.

Тоді прогнозний інтервал матиме вигляд:

$$\begin{aligned} \text{Н. М.: } ant \ln (\ln \hat{y}_{n+L} - S_y K^*), \\ \text{В. М.: } ant \ln (\ln \hat{y}_{n+L} + S_y K^*), \end{aligned}$$

$$\text{де } S_y = \sqrt{\frac{(\ln y_t - \ln \hat{y}_t)^2}{n-2}}.$$

5.2.5 Логарифмічна парабола

Розглядається тенденція виду $y_t = a b^t c^{t^2}$. Логарифмуємо цей вираз: $\ln y_t = \ln a + t \ln b + t^2 \ln c$. Це квадратична функція відносно параметра t . Оцінку параметрів a, b, c і прогнозний інтервал знаходимо так, як для многочлена другого ступеня.

Прогнозний інтервал буде таким:

$$\text{Н. М.: } \text{ant} \ln (\ln \hat{y}_{n+L} - S_y K^*),$$

$$\text{В. М.: } \text{ant} \ln (\ln \hat{y}_{n+L} + S_y K^*),$$

де K^* — табульовані значення для многочлена другого ступеня, а $S_y = \sqrt{\frac{(\ln y_t - \ln \hat{y}_t)^2}{n-3}}$.

5.2.6 Модифікована експонента

Тенденція має вигляд: $y_t = c - a b^t$. Вважаємо, що параметр $a > 0$. Вираз зведемо до лінійного виду: $\ln (c - y_t) = \ln a + t \ln b$. Позначимо $z_t = \ln (c^* - y_t)$. Вважаємо, що $c = c^*$ відоме, тобто відома асимптота. Отримаємо такий вираз для прогнозного інтервалу:

$$\text{Н. М.: } z_{n+L} - S_z K^*,$$

$$\text{В. М.: } z_{n+L} + S_z K^*,$$

де S_z — середньоквадратичне відхилення від тренду $z_t = \ln a + t \ln b$. Тоді прогнозний інтервал для y_{n+L} має вигляд

$$\text{Н. М.: } c^* - \text{ant} \ln (z_{n+L} - S_z K^*),$$

$$\text{В. М.: } c^* - \text{ant} \ln (z_{n+L} + S_z K^*).$$

Якщо параметр c невідомий, то можна оцінити параметри модифікованої експоненти $y = c + a b^t$ методом трьох сум. Згідно з цим методом часовий ряд розбиваємо на три однакових відрізка і позначимо через $\sum_1 y_t, \sum_2 y_t, \sum_3 y_t$ суми рівнів кожного з відрізків, n — кількість рівнів у кожному з відрізків. Використовуючи алгоритм

методу трьох сум, отримаємо такі оцінки параметрів c , a і b :

$$b = \sqrt[n]{\frac{\sum_3 y_t - \sum_2 y_t}{\sum_2 y_t - \sum_1 y_t}}, \quad a = \left(\sum_2 y_t - \sum_1 y_t \right) \frac{b - 1}{(b^n - 1)^2},$$

$$c = \frac{1}{n} \left[\frac{\sum_1 y_t \cdot \sum_3 y_t - (\sum_2 y_t)^2}{\sum_1 y_t + \sum_3 y_t - 2 \sum_2 y_t} \right].$$

Зауважимо, що метод трьох сум працює, якщо коливання ряду досить малі і результати не дуже чутливі до похибок. Тому перед оцінкою ряд необхідно згладити за допомогою ковзної середньої, якщо у ряді досить сильні коливання, або усунути досить великі викиди і замінити їх на усереднені.

5.2.7 Крива Гомперця

Тенденція, що описується кривою Гомперця має вигляд $y_t = c a^{b^t}$. За допомогою логарифмування криву Гомперця можна представити у вигляді модифікованої експоненти:

$$\ln y_t = \ln c + b^t \ln a.$$

Застосувавши метод трьох сум, отримаємо такі оцінки параметрів:

$$b = \sqrt[n]{\frac{\sum_3 \ln y_t - \sum_2 \ln y_t}{\sum_2 \ln y_t - \sum_1 \ln y_t}}, \quad \ln a = \left(\sum_2 \ln y_t - \sum_1 \ln y_t \right) \frac{b - 1}{(b^n - 1)^2},$$

$$c = \frac{1}{n} \left[\frac{\sum_1 \ln y_t \cdot \sum_3 \ln y_t - (\sum_2 \ln y_t)^2}{\sum_1 \ln y_t + \sum_3 \ln y_t - 2 \sum_2 \ln y_t} \right].$$

Слід пам'ятати про зауваження, що стосується розрахунку параметрів модифікованої експоненти відносно коливань і похибок.

5.2.8 Логістична крива Перла — Ріда

а) Якщо логістична крива має вигляд

$$\frac{1}{y_t} = k + a b^t,$$

то, застосувавши метод трьох сум, отримаємо такі оцінки параметрів:

$$b = \sqrt[n]{\frac{\sum_3 \frac{1}{y_t} - \sum_2 \frac{1}{y_t}}{\sum_2 \frac{1}{y_t} - \sum_1 \frac{1}{y_t}}}, \quad a = \left(\sum_2 \frac{1}{y_t} - \sum_1 \frac{1}{y_t} \right) \frac{b-1}{(b^n-1)^2},$$

$$k = \frac{1}{n} \left(\sum_1 \frac{1}{y_t} - \frac{b^n-1}{b-1} a \right) \quad \text{або} \quad k = \frac{1}{n} \left[\frac{\sum_1 \frac{1}{y_t} \cdot \sum_3 \frac{1}{y_t} - \left(\sum_2 \frac{1}{y_t} \right)^2}{\sum_1 \frac{1}{y_t} + \sum_3 \frac{1}{y_t} - 2 \sum_2 \frac{1}{y_t}} \right].$$

б) Нехай логістична крива представлена у вигляді

$$y_t = \frac{k}{1 + be^{-at}}.$$

Метод трьох сум дає такі оцінки параметрам логістичної кривої:

$$a = \frac{1}{n} (\ln D_1 - \ln D_2), \quad k = n : \left(\sum_1 \frac{1}{y_t} - \frac{D_1^2}{D_1 - D_2} \right),$$

$$b = \frac{k}{c} \frac{D_1^2}{D_1 - D_2},$$

$$\text{де } c = \frac{1 - e^{-na}}{1 - e^a}, \quad D_1 = \sum_1 \frac{1}{y_t} - \sum_2 \frac{1}{y_t}, \quad D_2 = \sum_2 \frac{1}{y_t} - \sum_3 \frac{1}{y_t}.$$

в) Якщо логістична крива має вигляд

$$y_t = \frac{k}{1 + 10^{a+bt}}$$

і відсутній повний ряд даних, то для оцінки параметрів можна скористатись методом трьох точок. Спочатку обираємо три точки, рівновіддалені одна від одної — на початку, середині й кінці ряду. Параметри визначаються з умови проходження логістичної кривої через ці три точки y_0 , y_1 і y_2 . Нехай відстань між y_0 і y_1 , y_1 і y_2 дорівнює n одиницям часу. Тоді отримаємо такі оцінки параметрів:

$$k = \frac{2y_0y_1y_2 - y_1^2(y_0 + y_2)}{y_0y_2 - y_1^2}, \quad b = \frac{1}{n} \log \frac{y_0(k - y_1)}{y_1(k - y_0)}, \quad a = \log \frac{k - y_0}{y_0}.$$

Так як при розрахунках використовується мало інформації про часовий ряд — всього три рівні і можливі похибки, то краще брати усереднені рівні, наприклад, середні геометричні.

Як зазначалося раніше, умовою застосування методу трьох точок є рівність відстаней по осі часу між обраними точками. Якщо ця умова не виконується,

то можна скористатися цим методом, але необхідно знати значення асимптоти. Знаючи різноманітну інформацію, наприклад, фізичні постійні, природні ресурси, границі споживання сировини і продуктів можна оцінити асимптоту. Тоді оцінку параметрів можна зробити за двома точками і значенням асимптоти $k = k^*$. Отримаємо такі оцінки параметрів:

$$a = \log \frac{k^* - y_0}{y_0}, \quad b = \frac{1}{n} \log \frac{y_0 (k^* - y_1)}{y_1 (k^* - y_0)}.$$

г) Таким же чином можна оцінити параметри логістичної кривої

$$y_t = \frac{k}{1 + be^{at}}.$$

Отримаємо такі оцінки параметрів:

$$a = \frac{1}{n} (\ln d_1 - \ln d_2), \quad k = 1 : \left(\frac{1}{y_0} - \frac{d_1^2}{d_1 - d_2} \right), \quad b = \frac{k - y_0}{y_0},$$

$$\text{де } d_1 = \frac{1}{y_0} - \frac{1}{y_1}, \quad d_2 = \frac{1}{y_1} - \frac{1}{y_2}.$$

Якщо значення k відомо, $k = k^*$, то параметри a і b можна оцінити так:

$$b = \frac{k^* - y_0}{y_0}, \quad a = \frac{1}{n} \ln \left(\frac{k^* y_1 - 1}{b} \right).$$

Метод трьох точок досить простий, але він чутливий до величин значень y_0, y_1, y_2 , які є, взагалі-то кажучи, випадковими. Але цей метод корисний особливо при уривчастих даних про досліджуване явище.

д) Розглянемо логістичну криву виду

$$y_t = \frac{k}{1 + ab^{-t}}.$$

Якщо нам відомо значення k , $k = k^*$, то зробимо такі перетворення:

$$\frac{y_t}{k^*} = \frac{1}{1 + ab^{-t}}, \quad \frac{k^*}{y_t} = 1 + ab^{-t}, \quad \frac{k^*}{y_t} - 1 = ab^{-t}.$$

$$\text{Далі } \ln \left(\frac{k^*}{y_t} - 1 \right) = \ln a - t \ln b.$$

Отримали лінійне рівняння відносно t . Далі для оцінки параметрів і прогнозного інтервалу використовуємо раніше розглянуту теорію.

5.3 Адаптивні моделі прогнозування

При прогнозуванні екстраполяційними методами вважається, що тенденція в минулому збережеться і в майбутньому.

Але якщо прогнозування відбувається в ситуації змінних умов або в коротко-строковому періоді, то найбільш важливими є останні рівні ряду. Тому оцінювати інформацію необхідно по-різному: більш пізню інформацію необхідно оцінювати з більшим коефіцієнтом оцінювання, ніж інформацію минулого. У цьому випадку найбільш ефективними методами прогнозування є адаптивні, що враховують нерівноцінність рівнів ряду.

Адаптивні методи прогнозування — це моделі дисконтування даних, які досить швидко пристосовують свої параметри до зміни умов.

Принцип дисконтування передбачає, що для побудови точних і надійних прогнозів більш пізніша інформація має більшу питому вагу по інформативності, ніж ранішня інформація.

При оцінці параметрів адаптивних моделей рівням ряду надають різні ваги в залежності від їхнього впливу на сучасний момент. Це дозволяє враховувати зміни в тенденції, а також коливання. Великі значення ваг, що належать проміжку $(0, 1)$, тобто більше 0.5, означає, що ці рівні більш цінні, а це останні рівні ряду. У попередніх рівнів ряду ваги менші 0.5. Перший випадок відповідає динамічним процесам, що швидко змінюються, а другий — більш стабільним.

Загальна схема побудови адаптивних моделей має такий вигляд. По декільком першим рівням ряду оцінюються параметри моделі. Далі будуємо прогноз за отриманою моделлю на один крок вперед. Відхилення прогнозу від фактичних рівнів є похибкою прогнозу, за допомогою якої робиться коригування моделі. По скоригованій моделі розраховується прогнозна оцінка на один крок і т. д. Таким чином, модель враховує нову інформацію і в кінці відображає тенденцію розвитку процесу в поточний період.

На принципі дисконтування інформації розроблені такі методи статистичного прогнозування:

- метод простого експоненціального згладжування;
- метод гармонічних ваг та інші.

Ці методи можуть бути використані при прогнозуванні тільки за таких умов:

- часові ряди повинні бути достатньо довгими;

- у рівнях часового ряду відсутні стрибки;
- зберігається принцип інерційності, тобто тенденції минулого і теперішнього можна продовжити на майбутнє.

5.3.1 Метод простого експоненціального згладжування

Він полягає у тому, що рівні початкового часового ряду згладжуються за допомогою ковзної середньої, ваги якої підпорядковуються експоненціальному закону розподілу. Ця ковзна середня отримала назву *експоненціальної середньої* — $S_t(y)$.

Особливість методу полягає в тому, що при розрахунку теоретичних значень за моделлю тренда враховується тільки значення попереднього рівня, взятого з певною вагою.

Загальна формула розрахунку експоненціальної середньої має вигляд:

$$S_t(y) = \alpha y_t + (1 - \alpha)S_{t-1}(y),$$

де $S_t(y)$ — значення експоненціальної середньої часового ряду для моменту t , $S_{t-1}(y)$ — значення експоненціальної середньої для моменту $(t - 1)$, y_t — значення останнього рівня ряду в момент часу t , α — параметр згладжування.

При розрахунку експоненціальної середньої $S_t(y)$ явно використовується значення тільки останнього рівня ряду.

У цьому методі дуже важливо знати, як обирати значення параметра α . Від значень α залежать ваги попередніх рівнів ряду і їх вплив на згладжуваний рівень.

Р. Г. Браун, автор метода, пропонує розраховувати параметр α за формулою: $\alpha = \frac{2}{n+1}$, де n — кількість рівнів, що увійшли в інтервал згладжування. На практиці параметр α обирається з проміжку (0.1; 0.3).

Параметр α можна визначити методом перебору різних його значень. При цьому в якості оптимального значення α обирається те значення α , при якому середня квадратична помилка прогнозу буде найменшою. Середньоквадратична помилка прогнозу розраховується за елементами всього ряду або за даними частини ряду, спеціально залишеної для перевірки якості прогнозової моделі. Це означає, що будується ретроспективний прогноз, сутність якого полягає в тому, що весь ряд розбивають на дві частини у співвідношенні 2/3 до 1/3. Для різних значень α будується модель прогнозу за першою частиною ряду (2/3), і за нею здійснюється прогноз на другу (1/3) частину ряду. За відхиленнями прогнозних значень ряду від емпіричних

розраховується середня квадратична похибка за формулою:

$$\sigma = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{n}}.$$

Крім знання параметра α необхідно знати $S_0(y)$. Тому існують два випадки:

- початкові умови (S_0) відомі;
- початкові умови невідомі.

Початкові значення відомі

У якості початкового значення S_0 можна використати середню арифметичну усіх значень рівнів ряду:

$$S_0 = \frac{\sum_{t=1}^n y_t}{n}.$$

Також за S_0 можна взяти y_1 .

Початкові значення невідомі

Розглянемо методи, що були розроблені Р. Г. Брауном.

Метод розрахунку початкових значень залежить від моделі тренда, що найкращим чином описує реальну тенденцію розвитку явища.

Якщо тенденція початкового часового ряду описується рівнянням лінійного тренду $\hat{y}_t = a_0 + a_1 t$, то алгоритм прогнозування методом простого експоненціального згладжування такий:

- 1) Методом найменших квадратів визначаються параметри лінійного тренду a_0 і a_1 : $\hat{y}_t = a_0 + a_1 t$.
- 2) Визначаються початкові умови першого і другого порядків (порядок початкових умов визначається кількістю параметрів рівняння тренду: лінійного тренду — a_0 і a_1) за формулами:

- початкова умова першого порядку:

$$S_0^1(y) = a_0 - \frac{1 - \alpha}{\alpha} a_1,$$

- початкова умова другого порядку:

$$S_0^2(y) = a_0 - \frac{2(1-\alpha)}{\alpha} a_1,$$

де a_0 і a_1 — параметри, визначені методом найменших квадратів.

3) Розраховуються експоненціальні середні першого і другого порядку:

- експоненціальна середня першого порядку:

$$S_t^1(y) = \alpha y_t + (1-\alpha)S_0^1(y),$$

де y_t — значення останнього фактичного рівня часового ряду;

- експоненціальна середня другого порядку:

$$S_t^2(y) = \alpha S_t^1(y) + (1-\alpha)S_0^2(y).$$

Далі розраховуємо параметри прогнозової моделі:

$$\begin{aligned}\hat{a}_0 &= 2S_t^1(y) - S_t^2(y), \\ \hat{a}_1 &= \frac{\alpha}{1-\alpha} [S_t^1(y) - S_t^2(y)].\end{aligned}$$

Прогноз здійснюємо за моделлю:

$$\hat{y}_{t+L} = \hat{a}_0 + \hat{a}_1 t.$$

4) Похибка прогнозу розраховується за формулою:

$$\sigma_{\hat{y}_{t+L}} = \sigma_y \sqrt{\frac{\alpha}{(2-\alpha)^5} [1 - 4(1-\alpha) + 5(1-\alpha)^2 + 2\alpha(4-3\alpha)] t + 2\alpha^2 L^2},$$

де σ_y — середня квадратична похибка, розрахована за відхиленнями емпіричних значень ознаки від теоретичних, отриманих за рівнянням тренду:

$$\sigma_y = \sqrt{\frac{\sum (y_t - \hat{y}_{t+L})^2}{n - k - 1}},$$

де k — кількість параметрів в рівнянні тренду.

Приклад 5.3. За даними

| | | | | | | | | | | | | |
|---------|------|------|------|------|------|------|------|------|------|------|------|------|
| Місяць | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Витрати | 21.4 | 22.1 | 23.9 | 24.3 | 24.9 | 26.9 | 28.0 | 28.5 | 28.8 | 28.6 | 29.3 | 31.9 |

побудувати прогноз витрат електроенергії населенням міста методом простого експоненціального згладжування, якщо тенденція описується лінійним рівнянням:

$$\hat{y}_t = 26.55 + 0.43t.$$

Визначаємо параметр α : $\alpha = \frac{2}{n+1} = \frac{2}{13} = 0.15$ ($n = 12$).

Далі,

$$S_0^1(y) = a_0 - \frac{1-\alpha}{\alpha}a_1 = 26.55 - \frac{1-0.15}{0.15} \cdot 0.43 = 24.11,$$

$$S_0^2(y) = a_0 - \frac{2(1-\alpha)}{\alpha}a_1 = 26.55 - \frac{2(1-0.15)}{0.15} \cdot 0.43 = 21.68,$$

$$S_t^1(y) = \alpha y_t + (1-\alpha)S_0^1(y) = 0.15 \cdot 31.9 + (1-0.15) \cdot 24.11 = 25.28,$$

$$S_t^2(y) = \alpha S_t^1(y) + (1-\alpha)S_0^2(y) = 0.15 \cdot 25.28 + (1-0.15) \cdot 21.68 = 22.22.$$

Тоді

$$\hat{a}_0 = 2S_t^1(y) - S_t^2(y) = 2 \cdot 25.28 - 22.22 = 28.34,$$

$$\hat{a}_1 = \frac{\alpha}{1-\alpha} [S_t^1(y) - S_t^2(y)] = \frac{0.15}{1-0.15} (25.28 - 22.22) = 0.54.$$

Отже, модель прогнозу витрат електроенергії населенням міста методом простого експоненціального згладжування має вигляд:

$$\hat{y}_{t+L} = 28.34 + 0.54t.$$



Якщо часовий ряд описується параболою другого порядку: $\hat{y}_t = a_0 + a_1t + a_2t^2$, то параметри визначаємо методом найменших квадратів і розраховуємо основні показники алгоритму.

Спочатку розрахуємо початкові умови:

– першого порядку

$$S_0^1(y) = a_0 - \frac{1-\alpha}{\alpha}a_1 + \frac{(1-\alpha)(2-\alpha)}{2\alpha^2}a_2,$$

– другого порядку

$$S_0^2(y) = a_0 - \frac{2(1-\alpha)}{\alpha}a_1 + \frac{(1-\alpha)(3-2\alpha)}{\alpha^2}a_2,$$

– третього порядку

$$S_0^3(y) = a_0 - \frac{3(1-\alpha)}{\alpha}a_1 + \frac{3(1-\alpha)(4-3\alpha)}{2\alpha^2}a_2.$$

Далі розраховуємо експоненціальні середні:

- першого порядку $S_t^1(y) = \alpha y_t + (1-\alpha)S_0^1(y)$,
- другого порядку $S_t^2(y) = \alpha S_t^1(y) + (1-\alpha)S_0^2(y)$,
- третього порядку $S_t^3(y) = \alpha S_t^2(y) + (1-\alpha)S_0^3(y)$.

Тоді параметри моделі прогнозу

$$\hat{y}_{t+L} = \hat{a}_0 + \hat{a}_1 t + \frac{1}{2}\hat{a}_2 t^2$$

оцінюються таким чином:

$$\begin{aligned}\hat{a}_0 &= 3 [S_t^1(y) - S_t^2(y)] + S_t^3(y), \\ \hat{a}_1 &= \frac{\alpha}{2(1-\alpha)^2} [(6-5\alpha)S_t^1(y) - 2(5-\alpha)S_t^2(y) + (4-3\alpha)S_t^3(y)], \\ \hat{a}_2 &= \frac{\alpha^2}{(1-\alpha)^2} [S_t^1(y) - 2S_t^2(y) + S_t^3(y)].\end{aligned}$$

Похибка прогнозу визначається за формулою:

$$\sigma_{\hat{y}_{t+L}} = \sigma_y \sqrt{2\alpha + 3\alpha^2 + 3\alpha^3 t},$$

де $\sigma_y = \sqrt{\frac{\sum (y_t - \hat{y}_{t+L})^2}{n - k - 1}}.$


Приклад 5.4. За даними попереднього прикладу побудувати модель прогнозу споживання електроенергії жителями міста методом простого експоненціального згладжування при умові, що тенденція зміни показника електроенергії описується рівнянням параболи другого порядку $\hat{y}_t = 26.86 + 0.43t - 0.007t^2$.

Маємо:

$$\begin{aligned}
 S_0^1(y) &= a_0 - \frac{1-\alpha}{\alpha} a_1 + \frac{(1-\alpha)(2-\alpha)}{2\alpha^2} a_2 = 26.86 - \frac{1-0.15}{0.15} \cdot 0.43 - \\
 &\quad - \frac{(1-0.15)(2-0.15)}{2 \cdot 0.15^2} \cdot 0.007 = 24.18, \\
 S_0^2(y) &= a_0 - \frac{2(1-\alpha)}{\alpha} a_1 + \frac{(1-\alpha)(3-2\alpha)}{2\alpha^2} a_2 = 26.86 - \frac{2(1-0.15)}{0.15} \cdot 0.43 - \\
 &\quad - \frac{(1-0.15)(3-2 \cdot 0.15)}{0.15^2} \cdot 0.007 = 21.276, \\
 S_0^3(y) &= a_0 - \frac{3(1-\alpha)}{\alpha} a_1 + \frac{3(1-\alpha)(4-3\alpha)}{2\alpha^2} a_2 = 26.86 - \frac{3(1-0.15)}{0.15} \cdot 0.43 + \\
 &\quad + \frac{3(1-0.15)(4-3 \cdot 0.15)}{2 \cdot 0.15^2} \cdot 0.007 = 18.14. \\
 S_t^1(y) &= \alpha y_t + (1-\alpha) S_0^1(y) = 0.15 \cdot 31.9 + (1-0.15) \cdot 24.18 = 25.34, \\
 S_t^2(y) &= \alpha S_t^1(y) + (1-\alpha) S_0^2(y) = 0.15 \cdot 25.34 + (1-0.15) \cdot 21.276 = 21.88, \\
 S_t^3(y) &= \alpha S_t^2(y) + (1-\alpha) S_0^3(y) = 0.15 \cdot 21.88 + (1-0.15) \cdot 18.14 = 18.701.
 \end{aligned}$$

Тоді

$$\begin{aligned}
 \hat{a}_0 &= 3 [S_t^1(y) - S_t^2(y)] + S_t^3(y) = 3(25.34 - 21.88) + 18.701 = 29.08, \\
 \hat{a}_1 &= \frac{\alpha}{2(1-\alpha)^2} [(6-5\alpha)S_t^1(y) - 2(5-4\alpha)S_t^2(y) + (4-3\alpha)S_t^3(y)] = \\
 &= \frac{0.15}{2(1-0.15)^2} [(6-5 \cdot 0.15) \cdot 25.34 - 2(5-4 \cdot 0.15) \cdot 21.88 + \\
 &\quad + (4-3 \cdot 0.15) \cdot 18.701] = 0.69, \\
 \hat{a}_2 &= \frac{\alpha^2}{(1-\alpha)^2} [S_t^1(y) - S_t^2(y) + S_t^3(y)] = \frac{0.15}{(1-0.15)^2} [25.34 - 2 \cdot \\
 &\quad \cdot 21.88 + 18.701] = 0.059.
 \end{aligned}$$

Таким чином, прогнозна модель споживання електроенергії жителями міста має вигляд: $\hat{y}_{t+L} = 29.08 + 0.69t + 0.059t^2$. 

5.4 Прогнозування рядів динаміки, що не мають тенденції

При розв'язуванні соціально-економічних задач інколи зустрічаються часові ряди без тенденції. Особливість прогнозування в таких рядах полягає в тому, що використовувати методи статистичного прогнозування для отримання точного або

інтервального прогнозу недоцільно. У цьому випадку доцільно застосовувати ймовірнісні статистичні методи прогнозного оцінювання.

Ймовірнісні методи оцінювання не дають можливості зробити точкову кількісну характеристику прогнозу. Вони дають оцінку ймовірності того, що значення прогнозу буде більше або менше значення останнього рівня початкового ряду. Ймовірнісні методи прогнозування дають менш точні прогнозні оцінки і мають невизначеність.

На практиці при аналізі часових рядів соціально-економічних явищ, що не мають тенденції, найбільше розповсюдження серед ймовірнісних методів прогнозування отримав метод, який ґрунтується на використанні закону розподілу Пуассона з густиною $\rho = e^\alpha$.

Розглянемо алгоритм методу.

- 1) Здійснюється послідовне порівняння кожного наступного значення рівня ряду зі значенням попереднього рівня. Якщо рівень збільшився, то ставимо знак «+», а при зменшенні ставимо «-». Перший рівень завжди мітимо знаком «-». Знак «+» показує, скільки періодів часу ряд зростає, і цей часовий період прийнято вважати сприятливою тенденцією. Цей метод завжди прогнозує сприятливу тенденцію.
- 2) Далі будуємо таблицю, що характеризує вид тенденції, довжину сприятливої тенденції (d) і частоту повторення сприятливої тенденції (c).

| Види тенденції | Довжина сприятливої тенденції d | Частота c |
|----------------|-----------------------------------|-------------|
| - - | 0 | |
| - + - | 1 | |
| - + + - | 2 | |
| - + + + - | 3 | |
| - + + + + - | 4 | |
| ... | ... | |

Записи робляться тільки у третьому стовпчику.

Довжина сприятливої тенденції (d) визначається кількістю плюсів між двома мінусами у ряді «+» і «-».

- 3) За даними таблиці визначається середня довжина сприятливої тенденції за формулою

$$\bar{d} = \frac{\sum dc}{\sum c},$$

де (d) — довжина сприятливої тенденції, а (c) — частота повторення сприятливої тенденції.

Середня довжина сприятливої тенденції показує, скільки у середньому в часовому ряді спостерігалася сприятлива тенденція.

Знаючи середню довжину сприятливої тенденції \bar{d} , визначаємо інтенсивність переривання цієї сприятливої тенденції (λ) за формулою $\lambda = \frac{1}{\bar{d}}$.

- 4) Ймовірність сприятливої тенденції визначається за законом розподілу Пуассона: $P = e^{-\lambda L}$, де

P — ймовірність того, що відбулася сприятлива тенденція,

λ — інтенсивність переривання сприятливої тенденції,

L — період випередження (кількість років збереження сприятливої тенденції).

Приклад 5.5. Є дані про споживання електроенергії жителями міста за дев'ять місяців:

| Місяць | Об'єм споживання | |
|----------|------------------|-----------------|
| | електроенергії | Знаки відхилень |
| | жителями міста | |
| січень | 4.5 | — |
| лютий | 5.2 | + |
| березень | 5.3 | + |
| квітень | 6.7 | + |
| травень | 6.1 | — |
| червень | 6.4 | + |
| липень | 5.8 | — |
| серпень | 5.0 | — |
| вересень | 5.3 | — |

Далі будемо по отриманим даним знаків відхилення ряд їх розподілу:

| Розподіл знаків відхилень | | |
|---------------------------|-------------------------------------|---------------|
| Вид тенденції | Довжина сприятливої тенденції d_i | Частота c_i |
| — — | 0 | 2 |
| — + — | 1 | 1 |
| — + + — | 2 | 0 |
| — + + + — | 3 | 1 |

Розрахуємо середню довжину сприятливої тенденції


$$\bar{d} = \frac{\sum c_i d_i}{\sum c_i} = \frac{0 \cdot 2 + 1 \cdot 1 + 2 \cdot 0 + 3 \cdot 1}{2 + 1 + 0 + 1} = 1.$$

Інтенсивність переривання сприятливої тенденції (λ) така:

$$\lambda = \frac{1}{\bar{d}} = \frac{1}{1} = 1.$$

Ймовірність спостереження сприятливої тенденції буде такою:

| Період збереження
сприятливої тенденції | t | λ | $-\lambda t$ | Ймовірність
сприятливої тенденції,
$e^{-\lambda t}$ |
|--|-----|-----------|--------------|---|
| жовтень | 1 | 1 | -1 | 0.368 |
| листопад | 2 | 1 | -2 | 0.135 |
| грудень | 3 | 1 | -3 | 0.049 |

Отримали, що ймовірність зростання споживання електроенергії жителями міста в жовтні в порівнянні з вереснем дорівнює 0.368. 

5.4.1 Оцінка точності прогнозу

Оцінка точності прогнозу є важливим етапом прогнозування. Для оцінки точності прогнозу використовується різниця між прогнозним значенням \hat{y}_t^* і фактичним y_t значенням показника. Цей підхід можна використовувати у таких випадках:

- період випередження відомий і відомі фактичні значення прогнозного показника;
- будується ретроспективний прогноз, тобто розраховуються прогнозні значення показника для періоду часу, для якого є фактичні значення.

У цьому випадку інформація ділиться на дві частини в співвідношенні 2/3 до 1/3. Перша частина значень рівнів використовується для визначення параметрів моделі прогнозу. Друга частина інформації служить для розрахунку оцінок прогнозу.

Розглянемо деякі показники точності прогнозу:

1) Абсолютна похибка прогнозу.

Вона визначається як різниця між емпіричними і прогнозними значеннями показника за формулою:

$$\Delta_a = y_t - \hat{y}_t^*,$$

де y_t — фактичні значення показника, \hat{y}_t^* — прогнознi значення показника.

2) Відносна похибка прогнозу.

Вона розраховується двома способами:

$$\Delta_b = \frac{\Delta_a}{y_t} = \frac{|y_t - \hat{y}_t^*|}{y_t} \cdot 100\%,$$

$$\Delta_b = \frac{\Delta_a}{\hat{y}_t} = \frac{|y_t - \hat{y}_t^*|}{\hat{y}_t} \cdot 100\%.$$

Зазначимо, що абсолютна і відносна похибки прогнозу є оцінкою точності одиничного прогнозу, що не дає можливості говорити про їх важливість в оцінці всієї прогнозної моделі. Тому на практиці інколи визначають не похибку прогнозу, а коефіцієнт якості прогнозу.

3) Коефіцієнт якості прогнозу.

Він визначається за формулою:

$$K = \frac{C}{C + H},$$

де C — кількість прогнозів, що співпали з фактичними значеннями, H — не співпали.

Якщо $K = 1$, то це означає, що всі значення прогнозних і фактичних значень співпадають і модель на 100 % описує явище.

4) Середній показник точності прогнозу.

Цей показник розраховується так:

$$\bar{\Delta} = \frac{\sum_{i=1}^n \Delta_t}{n} = \frac{\sum_{t=1}^n |y_t - \hat{y}_t|}{n},$$

де n — довжина частини або всього ряду, на якому порівнюються прогнознi і фактичні рівні. Показник показує узагальнену характеристику відхилень фактичних і прогнозних значень показника.

5) Середня квадратична похибка прогнозу.

Вона розраховується таким чином:

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

Між середньою абсолютною і середньою квадратичною похибками прогнозу існує таке співвідношення:

$$\sigma = 1.25\bar{\Delta}.$$

Недоліком двох останніх параметрів є їхня суттєва залежність від масштабу виміру рівнів.

6) Середня похибка апроксимації.

Щоб звільнитися від масштабу використовують середню похибку апроксимації

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t^*|}{y_t} 100\%.$$

Інтерпретація оцінки точності прогнозу за цим показником представлена в таблиці:

| <i>MAPE</i> | Інтерпретація точності |
|-------------|------------------------|
| < 10 | висока |
| 10 ÷ 20 | добра |
| 20 ÷ 50 | задовільна |
| > 50 | незадовільна |

7) Коефіцієнт невідповідності.

Цей показник був запропонований Г. Тейлом і має декілька модифікацій:

$$KH_1 = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t^* - y_t)^2}{\sum_{t=1}^n y_t^2}}, \quad KH_2 = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t^* - y_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2}},$$

де \bar{y} — середній рівень елементів ряду,

$$KH_3 = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t^* - y_t)^2}{\sum_{t=1}^n (y_t - \bar{y}_t)^2}},$$

де \bar{y}_t — теоретичні рівні часового ряду, отримані за моделлю тренда.

Приклади побудови моделей часового ряду

6.1 Алгоритм побудови моделі часового ряду

- 1) Згладжуємо часовий ряд ковзною середньою.
- 2) Оцінюємо сезонні коливання (S).
- 3) Будуємо аналітичний вигляд сезонної компоненти за допомогою ряду Фур'є.
- 4) Знаходимо значення тенденції із залишками.
- 5) Визначаємо аналітичний вигляд тенденції (T).
- 6) Розраховуємо аналітичний вигляд тенденції.
- 7) Знаходимо суму ($T + S$).
- 8) Розраховуємо залишки.
- 9) Перевіряємо модель на адекватність.
- 10) Робимо прогноз за отриманою моделлю.
- 11) Оцінюємо точність прогнозу.

Якщо відсутні сезонні коливання, то при побудові моделі пункти 2, 3 і 4 виключаємо з розгляду.

Приклад 6.1 (Побудова адитивної моделі часового ряду). Розглянемо дані про об'єм споживання електроенергії жителями міста за чотири роки (16 кварталів).

Раніше встановлено, що часовий ряд має сезонні коливання з періодичністю 4. Дивлячись на графік можемо зробити висновок, що амплітуда коливань майже однакова. Це означає, що часовий ряд має адитивну модель. Далі розрахуємо складові моделі за вказаним раніше алгоритмом.

| № квар-
тала | Споживання
електроенергії | Всього
за чотири
квартали | Ковзна середня
за 4 квартали | Центрована
ковзна
середня | Оцінка
сезонної
компоненти |
|-----------------|------------------------------|---------------------------------|---------------------------------|---------------------------------|----------------------------------|
| 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 | 39.6 | 7.40 | 7.300 | 2.700 |
| 9 | 8.0 | 30.0 | 7.50 | 7.450 | 0.550 |
| 10 | 5.6 | 31.0 | 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 | – | – | – | – |

Табл. 6.1. Вихідні дані про об'єм споживання електроенергії

1) Проведемо вирівнювання рівнів початкового ряду методом ковзної середньої.
Для цього:

- а) складемо рівні ряду послідовно за кожні чотири квартали зі зсувом на один квартал і визначимо умовні об'єми споживання електроенергії за рік (графа 3);
- б) поділимо отримані суми на 4 і знайдемо ковзні середні (графа 4). Зазначимо, що вирівнені значення не містять у собі сезонної компоненти;
- в) приведемо отримані значення у відповідність з фактичними моментами часу, для цього знаходимо центровані ковзні середні — середні арифметичні двох послідовних значень (графа 5).

2) Знайдемо оцінки сезонної компоненти як різницю між фактичними рівнями ряду і центральними ковзними середніми (графа 6). Використаємо ці оцінки для розрахунку значень сезонної компоненти S . Для цього знайдемо середні за кожний квартал (за всі роки) оцінки сезонної компоненти \bar{S}_t . Якщо у моделі присутня сезонна компонента, то сезонні коливання за період взаємознищуються. Для адитивної моделі сума значень сезонної компоненти по всім кварталам дорівнює нулеві.

Маємо $0.6 - 1.958 - 1.275 + 2.708 = 0.075$. Визначимо коригуючий коефіцієнт:

$$k = 0.075 : 4 = 0.01875.$$

| Показники | Рік | № квартала | | | |
|---|-----|------------|--------|--------|-------|
| | | I | II | III | IV |
| | 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 | – | – |
| Всього за i -й квартал
(за всі роки) | | 1.800 | –5.875 | –3.825 | 8.125 |
| Середня оцінка сезонної
компоненти для i -го квартала, \bar{S}_i | | 0.600 | –1.958 | –1.275 | 2.708 |
| Скоригована сезонна
компонента S_t | | 0.581 | –1.977 | –1.294 | 2.690 |

Табл. 6.2. Розрахунок значень сезонної компоненти

Розрахуємо скориговані значення сезонної компоненти як різницю між її середньою оцінкою і коригуючим коефіцієнтом k : $S_t = \bar{S}_t - k$, $t = 1, 2, 3, 4$.

Перевіримо умову рівності нулевій суми значень скоригованих сезонних компонент: $0.581 - 1.977 - 1.294 + 2.690 = 0$.

Отже, отримали такі значення сезонної компоненти:

$$\begin{aligned} \text{I квартал: } S_1 &= 0.581 & \text{II квартал: } S_2 &= -1.977 \\ \text{III квартал: } S_3 &= -1.294 & \text{IV квартал: } S_4 &= 2.690 \end{aligned}$$

Занесемо отримані значення у [таблицю](#) у графу 3.

- 3) Знайдемо аналітичний вигляд сезонної компоненти за допомогою ряду Фур'є. Кількість гармонік $M = 4$. Скористаємося формулами і отримаємо такі коефіцієнти:

$$\begin{aligned} a_0 &= 0, & a_1 &= a_2 = a_3 = 0, & a_4 &= 0.9375; \\ b_1 &= b_2 = b_3 = 0, & b_4 &= -2.3335. \end{aligned}$$

Отже, аналітичний вигляд сезонної компоненти наступний:

$$\hat{S} = 0.0375 \cos \frac{\pi}{2}x - 2.3335 \sin \frac{\pi}{2}x.$$

Розрахуємо за цією формулою значення сезонної компоненти у всіх точках ряду. Значення занесемо у [графу 4](#).

- 4) Вирахуємо значення сезонної компоненти з кожного рівня початкового часового ряду ([графа 5](#)).

| t | y_t | S_t | \hat{S}_t | $T_t + \varepsilon_t = y_t - \hat{S}_t$ | T | $T + \hat{S}$ | $\varepsilon_t = y_t - (T_t + \hat{S})$ |
|-----|-------|--------|-------------|---|-------|---------------|---|
| 1 | 6.0 | 0.581 | 0.9375 | 5.063 | 5.839 | 6.777 | -0.777 |
| 2 | 4.4 | -1.977 | -2.3335 | 6.734 | 6.034 | 3.700 | 0.700 |
| 3 | 5.0 | -1.294 | -0.9375 | 5.938 | 6.229 | 5.291 | -0.291 |
| 4 | 9.0 | 2.690 | 2.3335 | 6.667 | 6.424 | 8.758 | 0.242 |
| 5 | 7.2 | 0.581 | 0.9375 | 6.263 | 6.619 | 7.557 | -0.357 |
| 6 | 4.8 | -1.977 | -2.3335 | 7.134 | 6.814 | 4.480 | 0.320 |
| 7 | 6.0 | -1.294 | -0.9375 | 6.938 | 7.009 | 6.071 | -0.071 |
| 8 | 10.0 | 2.690 | 2.3335 | 7.667 | 7.204 | 9.538 | 0.462 |
| 9 | 8.0 | 0.581 | 0.9375 | 7.063 | 7.399 | 8.337 | -0.337 |
| 10 | 5.6 | -1.977 | -2.3335 | 7.934 | 7.594 | 5.260 | 0.340 |
| 11 | 6.4 | -1.294 | -0.9375 | 7.338 | 7.789 | 6.851 | -0.451 |
| 12 | 11.0 | 2.690 | 2.3335 | 8.667 | 7.984 | 10.318 | 0.682 |
| 13 | 9.0 | 0.581 | 0.9375 | 8.063 | 8.179 | 9.117 | -0.117 |
| 14 | 6.6 | -1.977 | -2.3335 | 8.934 | 8.374 | 6.040 | 0.560 |
| 15 | 7.0 | -1.294 | -0.9375 | 7.938 | 8.569 | 7.631 | -0.631 |
| 16 | 10.8 | 2.690 | 2.3335 | 8.467 | 8.764 | 11.098 | -0.298 |

Табл. 6.3. Розрахунок вирівнених значень T і помилок ε

- 5) Визначимо компоненту T . Для цього скористаємося методом найменших квадратів. Для розрахунків скористаємося значеннями [графі 5](#). Отримаємо, що $T = 5.644 + 0.195t$. Далі розраховуємо значення рівнів T для $t = \overline{1, 16}$. Отримані значення занесемо в [графу 6](#).
- 6) Далі знайдемо суму компонент T і \hat{S} . Значення $T + \hat{S}$ занесемо в [графу 7](#).
- 7) Розрахуємо залишки $\varepsilon_t = y_t - (T_t + \hat{S}_t)$. Їх значення заносимо в [графу 8](#).
- 8) Перевіряємо модель на адекватність. Спочатку встановимо, чи підпорядковуються залишки нормальному розподілу. Скористаємося критерієм Девіда — Харті — Пірсона:

$$U = \frac{R}{S} = \frac{\varepsilon_{\max} - \varepsilon_{\min}}{\sigma_{\varepsilon}} = 3.205.$$

Нехай рівень значущості $\alpha = 0.05$. Тоді згідно [таблиці](#) $U_1 = 3.01$, $U_2 = 4.24$. Так як $U_1 < U < U_2$, то приймаємо гіпотезу про нормальність залишків.

- 9) Перевіримо гіпотезу про рівність нулю математичного очікування залишків. Для цього розрахуємо

$$t_{\text{СТ}} = \frac{\bar{\varepsilon}}{\sigma_{\varepsilon}} \sqrt{n} = -0.013.$$

Так як $|t_{\text{СТ}}| < t_{\text{КР}} = t(0.05; 16 - 1) = 2.1314$, то приймаємо гіпотезу про рівність математичного очікування залишків нулю.

- 10) Оцінюємо точність моделі. Спочатку розрахуємо суму квадратів залишків

(RSS):


$$RSS = \sum_{t=1}^{16} \varepsilon_t^2 = 3.397.$$

Визначимо середню відносну похибку апроксимації ($MAPE$):

$$MAPE = \frac{100\%}{16} \sum_{t=1}^{16} \frac{|\varepsilon_t|}{y_t} = 6.243\%.$$

Робимо висновок, що модель має високу точність апроксимації. Тепер обчислимо коефіцієнт детермінації (R^2):

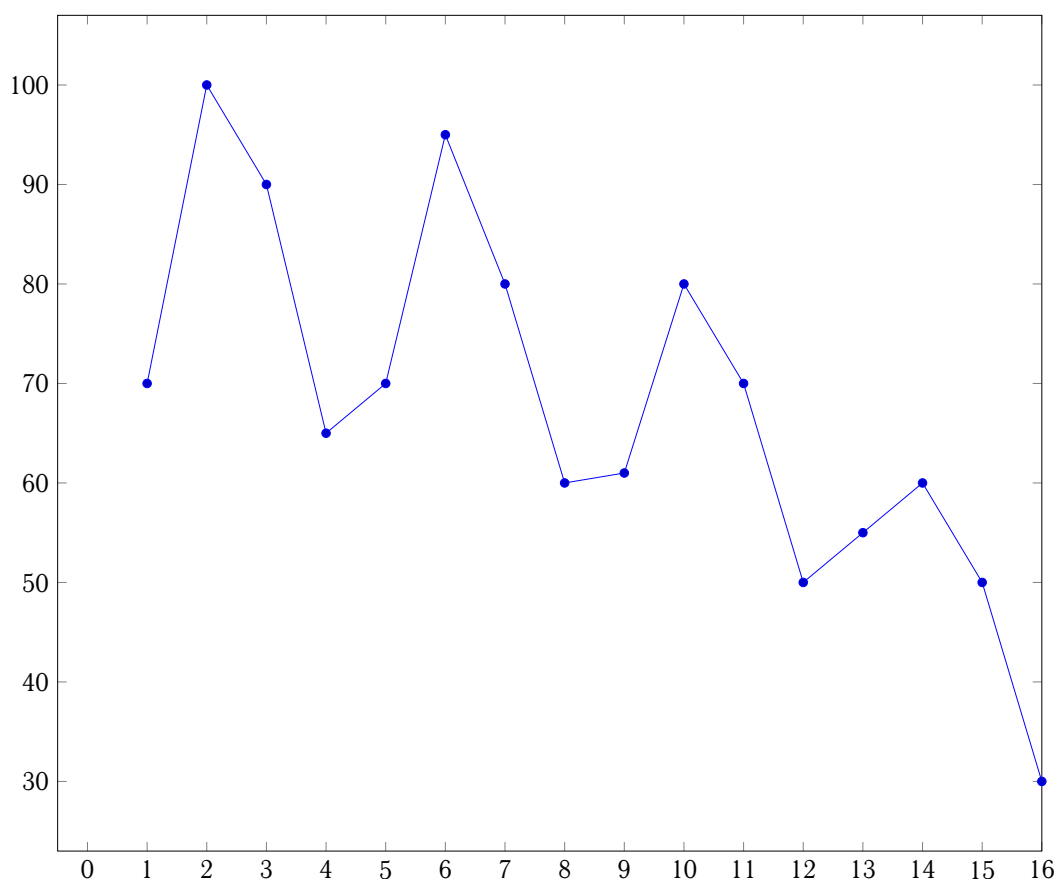
$$R^2 = 1 - \frac{RSS}{\sum_{t=1}^{16} (y_t - \bar{y})^2} = 0.949.$$

Коефіцієнт детермінації близький до одиниці, що також свідчить про високу точність моделі. 

Приклад 6.2 (побудова мультиплікативної моделі часового ряду). Нехай маємо поквартальні дані про прибуток компанії за чотири роки.

| Рік \ Квартал | Квартал | | | |
|---------------|---------|-----|-----|----|
| | I | II | III | IV |
| 1 | 72 | 100 | 90 | 64 |
| 2 | 70 | 92 | 80 | 58 |
| 3 | 62 | 80 | 68 | 48 |
| 4 | 52 | 60 | 50 | 30 |

Графік часового ряду вказує на наявність сезонних коливань з періодом в чотири квартали і з спадною тенденцією рівнів ряду. Так як амплітуда сезонних коливань спадає, то вважаємо, що існує мультиплікативна залежність між компонентами ряду, тобто $y = T \cdot S \cdot \varepsilon$.



- 1) Зробимо вирівнювання початкових рівнів ряду методом ковзної середньої. Все робимо так, як у попередньому прикладі.
- 2) Знайдемо оцінки сезонної компоненти як результат ділення фактичних рівнів ряду на центровані ковзні середні (графа 6). Використаємо ці оцінки для розрахунку значень сезонної компоненти S . Для цього знайдемо середні за кожний

квартал оцінки сезонної компоненти S_t . У мультиплікативній моделі сума значень сезонної компоненти по всіх кварталах повинна дорівнювати кількості періодів у циклі. У нашому прикладі кількість періодів одного циклу (року) дорівнює 4 (чотири квартали).

| № квартала | Прибуток компанії y_t | Всього за 4 квартала | Ковзна середня за 4 квартала | Центрована ковзна середня | Оцінка сезонної компоненти |
|------------|-------------------------|----------------------|------------------------------|---------------------------|----------------------------|
| 1 | 72 | — | — | — | — |
| 2 | 100 | 326 | 81.5 | — | — |
| 3 | 90 | 324 | 81.0 | 81.25 | 1.108 |
| 4 | 64 | 316 | 79.0 | 80.00 | 0.800 |
| 5 | 70 | 306 | 76.5 | 77.75 | 0.900 |
| 6 | 92 | 300 | 75.0 | 75.75 | 1.215 |
| 7 | 80 | 292 | 73.0 | 74.00 | 1.081 |
| 8 | 58 | 280 | 70.0 | 71.50 | 0.811 |
| 9 | 62 | 268 | 67.0 | 68.50 | 0.905 |
| 10 | 80 | 258 | 64.5 | 65.75 | 1.217 |
| 11 | 68 | 248 | 62.0 | 63.25 | 1.075 |
| 12 | 48 | 228 | 57.0 | 59.50 | 0.807 |
| 13 | 52 | 210 | 52.5 | 54.75 | 0.950 |
| 14 | 60 | 192 | 48.0 | 50.25 | 1.194 |
| 15 | 50 | — | — | — | — |
| 16 | 30 | — | — | — | — |

Табл. 6.4. Розрахунок оцінок сезонної компоненти

| Показники | Рік | № квартала | | | |
|--|-----|------------|-------|-------|-------|
| | | I | II | III | IV |
| | 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 | — | — |
| Всього за t -й квартал (за всі роки) | | 2.755 | 3.626 | 3.264 | 2.424 |
| Середня оцінка сезонної компоненти для t -го квартала, \bar{S}_t | | 0.918 | 1.209 | 1.088 | 0.808 |
| Скоригована сезонна компонента S_t | | 0.913 | 1.202 | 1.082 | 0.803 |

Табл. 6.5. Розрахунок сезонної компоненти

Далі маємо $0.918 + 1.209 + 1.088 + 0.808 = 4.023$.

Визначаємо коригуючий коефіцієнт: $k = 4/4.023 = 0.9943$.

Скориговані значення сезонної компоненти отримаємо таким чином: $S_t = \bar{S}_t \cdot k$, $t = 1, 2, 3, 4$.

Тепер знайдемо суму значень сезонної компоненти і перевіримо, чи сума дорівнює 4: $0.913 + 1.202 + 1.082 + 0.803 = 4$.

Отже, отримали такі значення сезонної компоненти:

$$\begin{array}{cccc} \text{I квартал} & \text{II квартал} & \text{III квартал} & \text{IV квартал} \\ S_1 = 0.913 & S_2 = 1.202 & S_3 = 1.082 & S_4 = 0.803 \end{array}$$

Занесемо отримані значення S_t в таблицю.

- 3) Поділимо кожен рівень початкового ряду на відповідні значення сезонної компоненти. Таким чином, отримаємо величини $T \cdot E = Y : S$ (графа 4), що включають тільки тенденцію і випадкову величину.

| t | y_t | S_t | $T \cdot \varepsilon = Y : S$ | T | $T \cdot S$ | $\varepsilon = Y : (T \cdot S)$ | $\varepsilon^1 = Y - (T \cdot S)$ | $(\varepsilon^1)^2$ |
|-----|-------|-------|-------------------------------|-------|-------------|---------------------------------|-----------------------------------|---------------------|
| 1 | 72 | 0.913 | 78.86 | 87.80 | 80.16 | 0.898 | -8.16 | 66.66 |
| 2 | 100 | 1.202 | 83.19 | 85.03 | 102.20 | 0.978 | -2.20 | 4.86 |
| 3 | 90 | 1.082 | 83.18 | 82.25 | 89.00 | 1.011 | 1.00 | 1.00 |
| 4 | 64 | 0.803 | 79.70 | 79.48 | 63.82 | 1.003 | 0.18 | 0.03 |
| 5 | 70 | 0.913 | 76.67 | 76.70 | 70.03 | 1.00 | -0.03 | 0.00 |
| 6 | 92 | 1.202 | 76.54 | 73.93 | 88.86 | 1.035 | 3.14 | 9.85 |
| 7 | 80 | 1.082 | 73.94 | 71.15 | 16.99 | 1.039 | 3.01 | 9.08 |
| 8 | 58 | 0.803 | 72.23 | 68.38 | 4.91 | 1.056 | 3.09 | 9.57 |
| 9 | 62 | 0.913 | 67.91 | 65.60 | 59.90 | 1.035 | 2.10 | 4.43 |
| 10 | 80 | 1.202 | 66.56 | 62.83 | 75.52 | 1.059 | 4.48 | 20.08 |
| 11 | 68 | 1.082 | 62.85 | 60.05 | 64.98 | 1.047 | 3.02 | 9.14 |
| 12 | 48 | 0.803 | 59.78 | 57.28 | 45.99 | 1.044 | 2.01 | 4.03 |
| 13 | 52 | 0.913 | 56.96 | 54.50 | 49.76 | 1.045 | 2.24 | 5.02 |
| 14 | 60 | 1.202 | 49.92 | 51.73 | 62.18 | 0.965 | -2.18 | 4.73 |
| 15 | 50 | 1.082 | 46.21 | 48.95 | 52.97 | 0.944 | -2.97 | 8.79 |
| 16 | 30 | 0.803 | 37.36 | 46.18 | 37.08 | 0.809 | -7.08 | 50.12 |


Табл. 6.6. Розрахунок вирівнених значень T і похибок ε

- 4) Розрахуємо компоненту T . Для цього розрахуємо параметри лінійного тренду, використовуючи рівні $(T \cdot \varepsilon)$. Метод найменших квадратів дає таку модель:

$$T = 90.59 - 2.773t.$$

В отримане рівняння підкладемо $t = \overline{1, 16}$ і отримаємо рівні T для всіх моментів часу (графа 5).

- 5) Розрахуємо значення $T \cdot S$ (графа 6).

- 6) Далі визначаємо похибку в мультиплікативній моделі: $\varepsilon = Y : (T \cdot S)$ (графа 7). Для порівняння мультиплікативної моделі з іншими моделями часового ряду по аналогії з адитивною моделлю використовують суму квадратів абсолютних похибок. 

Застосування фіктивних змінних для моделювання сезонних коливань

Моделювати часовий ряд з сезонними коливаннями можна за допомогою побудови регресії з включенням фактора часу і фіктивних змінних. Кількість фіктивних змінних в такій моделі повинна бути на одиницю менше кількості періодів часу всередині одного циклу коливань. Наприклад, якщо моделюємо поквартальні дані, то модель включатиме фактор часу і три фіктивні змінні. Кожна фіктивна змінна відображає сезонну (циклічну) компоненту часового ряду для якогось одного періоду. Вона дорівнює одиниці для одного періоду і нулеві для інших.

Якщо розглядається часовий ряд з циклічними коливаннями з періодичністю k , то модель з фіктивними змінними матиме вигляд:

$$y_t = a + bt + c_1x_1 + \dots + c_{k-1}x_{k-1} + \varepsilon_t,$$

$$\text{де } x_j = \begin{cases} 1, & \text{для кожного } j \text{ всередині кожного циклу,} \\ 0, & \text{в інших випадках.} \end{cases}$$

Наприклад, якщо моделювати сезонні коливання на основі поквартальних даних за декілька років, кількість кварталів в одному році $k = 4$, а модель матиме вигляд:

$$y_t = a + bt + c_1x_1 + c_2x_2 + c_3x_3 + \varepsilon_t.$$

Тоді рівняння тренду для кожного квартала матиме такий вигляд:

для 1 квартала: $y_t = a + bt + c_1 + \varepsilon_t$

для 2 квартала: $y_t = a + bt + c_2 + \varepsilon_t$

для 3 квартала: $y_t = a + bt + c_3 + \varepsilon_t$

для 4 квартала: $y_t = a + bt + \varepsilon_t$

Це рівняння є аналог адитивної моделі часового ряду, так як фактичний рівень часового ряду є сума трендової, сезонної і випадкової компоненти.

Приклад 7.1 (побудова моделі регресії часового ряду з фіктивними змінними). По даним прикладу про споживання електроенергії побудуємо модель регресії з фіктивними змінними.

Модель матиме вигляд: $y_t = a + bt + c_1x_1 + c_2x_2 + c_3x_3 + \varepsilon_t$.

Матриця початкових даних матиме вигляд:

| y | t | x_1 | x_2 | x_3 |
|------|-----|-------|-------|-------|
| 6.0 | 1 | 1 | 0 | 0 |
| 4.4 | 2 | 0 | 1 | 0 |
| 5.0 | 3 | 0 | 0 | 1 |
| 9.0 | 4 | 0 | 0 | 0 |
| 7.2 | 5 | 1 | 0 | 0 |
| 4.8 | 6 | 0 | 1 | 0 |
| 6.0 | 7 | 0 | 0 | 1 |
| 10.0 | 8 | 0 | 0 | 0 |
| 8.0 | 9 | 1 | 0 | 0 |
| 5.6 | 10 | 0 | 1 | 0 |
| 6.4 | 11 | 0 | 0 | 1 |
| 11.0 | 12 | 0 | 0 | 0 |
| 9.0 | 13 | 1 | 0 | 0 |
| 6.6 | 14 | 0 | 1 | 0 |
| 7.0 | 15 | 0 | 0 | 1 |
| 10.8 | 16 | 0 | 0 | 0 |

Методом найменших квадратів отримаємо таку регресію: $\hat{y}_t = 8.3250 + 0.1875t - 2.0875x_1 - 4.0475x_2 - 3.9125x_3$.

Проаналізуємо отримані результати. Коефіцієнт детермінації моделі $R^2 = 0.985$ досить високий, близький до одиниці. Параметр $a = 8.3250$ є сума початкового рівня ряду і сезонної компоненти в IV кварталі (так як для четвертого кварталу $x_1 = x_2 = x_3 = 0$). Сезонні коливання в I, II, III кварталах зменшують цю величину (від'ємні знаки при x_1, x_2, x_3). Зазначимо, що ці параметри при x_i не дорівнюють

значенням сезонної компоненти, так як вони характеризують не сезонні зміни рівнів ряду, а їх відхилення від рівнів, що враховують сезонні впливи в IV кварталі. Додатна величина $b = 0.1875$ вказує про присутність зростаючого тренду.

Основний недолік моделі з фіктивними змінними той, що до моделі може вводитися велика кількість фіктивних змінних, а тому довжину ряду необхідно брати досить великою. ☒

7.1 Моделювання тенденції часового ряду при наявності структурних змін

Розглянемо випадок лінійної тенденції. Існують одночасні зміни характеру тенденції часового ряду, що викликані структурними змінами в аналізованому процесі. Тобто починаючи з деякого моменту часу t^* відбувається зміна характеру динаміки ряду, див. [рис. 7.1](#).

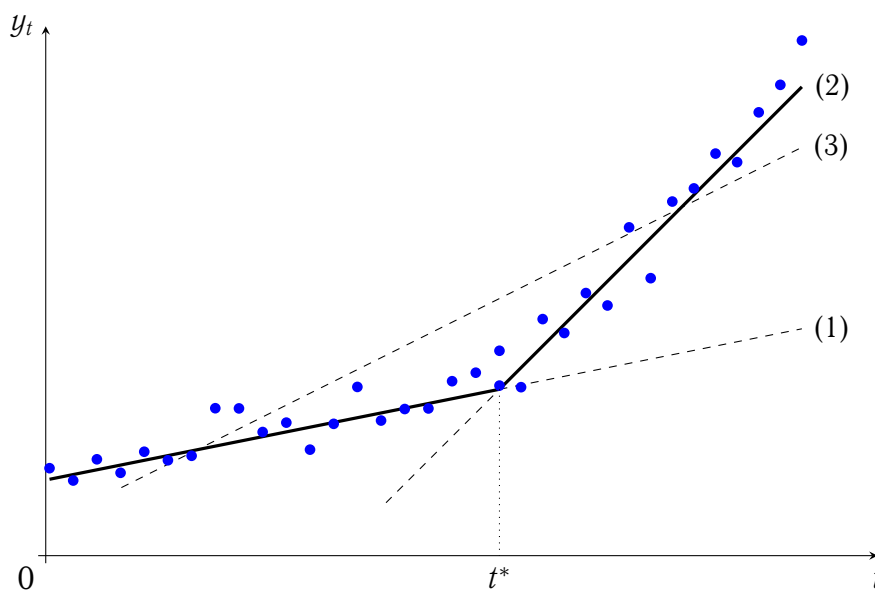


Рис. 7.1. Часовий ряд зі структурними змінами

У момент t^* відбуваються значні зміни факторів, що входять до часового ряду. Тому виникає питання, чи значимо впливають зміни факторів на характер тенденції.

Якщо вплив цей значущий, то для моделювання тенденції даного часового ряду необхідно використати кусково-лінійні моделі регресії. Це означає, що ряд розбиваємо на дві частини: одна до моменту t^* , а друга — після.

Далі по кожній із частин будемо рівняння лінійної регресії (1) і (2).

Коли структурні зміни незначимо впливають на характер часового ряду, то будуємо одну регресію (3).

Для визначення типу регресії необхідно скористатися тестом Г. Чоу. Для цього введемо деякі позначення і сформуємо з них таблицю. Цією таблицею можна користуватися для випадку трендів, відмінних від лінійного.

| № рівняння | Вид рівняння | Кількість спостережень | Залишкова сума квадратів | Кількість параметрів у рівнянні | Кількість ступенів свободи залишкової дисперсії |
|---|--------------|------------------------|--------------------------|---------------------------------|---|
| Модель з декількох рівнянь | | | | | |
| (1) | y^1 | n_1 | S^1 | p_1 | $n_1 - p_1$ |
| (2) | y^2 | n_2 | S^2 | p_2 | $n_2 - p_2$ |
| Рівняння тренду за всією сукупністю даних | | | | | |
| (3) | y^3 | n | S^3 | p_3 | $n - p_3 = (n_1 + n_2) - p_3$ |

Якщо розглядається лінійна модель, то $p_1 = p_2 = p_3 = 2$.

Залишкова сума квадратів по кусково-лінійній моделі дорівнює $S = S^1 + S^2$, а відповідна кількість ступенів свободи є $(n_1 - p_1) + (n_2 - p_2) = (n - p_1 - p_2)$.

При переході від одного рівняння тренду до кусково-лінійної моделі скорочення залишкової суми квадратів таке: $\Delta S = S^3 - S$.

Кількість ступенів свободи, що відповідає ΔS , дорівнює $n - p_3 = (n - p_1 - p_2) = p_1 + p_2 - p_3$.

Розрахуємо F_{CT} :

$$F_{CT} = \frac{\Delta S (n - p_1 - p_2)}{S (p_1 + p_2 - p_3)}.$$

Якщо $F_{CT} > F_{KP} = F(\alpha; p_1 + p_2 - p_3; n - p_1 - p_2)$, то з ймовірністю $(1 - \alpha) 100\%$ гіпотеза про структурну стабільність тенденції відхиляється, а вплив структурних змін на динаміку ряду вважають значущим. Це означає, що необхідно будувати кусково-лінійну модель.

Якщо $F_{CT} < F_{KP}$, то модель структурно стабільна, тобто рівняння (1) і (2) описують одну і ту саму тенденцію, і тому відмінність коефіцієнтів цих рівнянь статистично незначуща.

Далі розглянемо випадки структурної нестабільності тенденції. Для лінійної тенденції $y^1 = a_1 + b_1 t$, $y^2 = a_2 + b_2 t$.

- 1) Якщо a_1 і a_2 значимо відрізняються між собою, а b_1 і b_2 — ні, то геометрично це означає, що прямі y^1 і y^2 паралельні. Тобто відбулася стрибкоподібна зміна рівнів ряду y_t у момент часу t^* при незмінному середньому абсолютному прирості (рис. 7.2a).
- 2) Параметри a_1 і a_2 статистично незначимо відрізняються, а b_1 і b_2 — статистично значимо різняться. Це означає, що прямі y^1 і y^2 перетинають вісь ординат в одній точці, а зміна тенденції пов'язана зі зміною середнього абсолютного приросту у момент t^* (рис. 7.2b).
- 3) Значимо відрізняються параметри a_1 і a_2 , а також b_1 і b_2 . У цьому випадку початкові рівні різні і середні абсолютні прирости (рис. 7.2c).

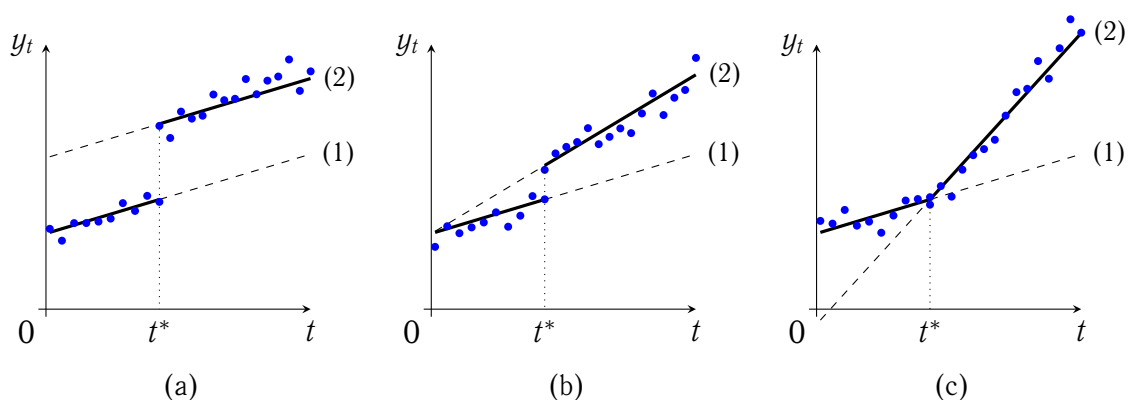


Рис. 7.2. Структурна нестабільність лінійної тенденції

Розглянемо підхід Д. Гуйараті дослідження тенденції для розглянутого вище випадку. Д. Гуйараті включає до моделі регресії фіктивну змінну z_t , яка приймає значення 1 для всіх $t < t^*$ і 0 для всіх $t > t^*$. Далі розглядається модель

$$y_t = a + b \cdot z_t + c \cdot t + d \cdot (z_t \cdot t) + \varepsilon_t. \quad (*)$$

Для кожного з проміжків отримаємо такі рівняння:

$$\begin{aligned} \text{для } t < t^*: \quad z_t &= 1, \quad y_t = (a + b) + (c + d)t + \varepsilon_t, \\ \text{для } t > t^*: \quad z_t &= 0, \quad y_t = a + ct + \varepsilon_t. \end{aligned}$$

Якщо порівняти рівняння y^1 , y^2 із (*), то отримаємо, що

$$a_1 = (a + b), \quad b_1 = (c + d), \quad a_2 = a, \quad b_2 = c.$$

Отже, $b = a_1 - a_2$, $d = b_1 - b_2$. Тоді оцінка статистичної значущості відмінності a_1 і a_2 , а також b_1 і b_2 еквівалентна статистичній значущості параметрів b і d рівняння (*).

Якщо у рівнянні (*) b є статистично значущим, а d — ні, то маємо випадок 1) і [рис. 7.2а](#). Якщо ж d є статистично значущим, а b — ні, то маємо випадок 2), [рис. 7.2б](#). І якщо b і d статистично значущі, то отримали випадок 3), [рис. 7.2с](#).

Підхід Д. Гуйараті простіший, так як використовується для дослідження тільки одне рівняння.

ДОДАТОК А

Завдання для самостійної роботи

Побудувати модель часового ряду за наступною схемою:

1. Побудувати графік часового ряду.
2. Обрати клас моделі — адитивна чи мультиплікативна. Обґрунтувати свій вибір.
3. Провести вирівнювання за допомогою ковзної середньої. Оцінити чисельно сезонну компоненту ряду (S).
4. Знайти аналітичний вигляд сезонної компоненти за допомогою ряду Фур'є.
5. Оцінити чисельно тенденцію ряду (T).
6. Визначити аналітичний вигляд тенденції.
7. Розрахувати залишки.
8. Перевірити адекватність моделі.
9. Оцінити точність моделі за допомогою показників $MAPE$ та R^2 .
10. Зробити прогноз за отриманою моделлю та оцінити точність прогнозу.

Початкові дані

Задано часовий ряд y_t , $t \in \overline{1, 50}$.

Варіанти 1–10

| t | Варіант | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----------|---------|-------|--------|-------|-------|-------|--------|-------|-------|-------|-------|
| | | | | | | | | | | | |
| 1 | | 2.17 | 3.17 | 3.17 | 2.17 | 2.17 | 1.88 | 0.88 | 0.88 | 0.88 | 1.88 |
| 2 | | 5.32 | 40.42 | 12.11 | 10.99 | 11.23 | 12.15 | 6.59 | 2.70 | 4.69 | 7.64 |
| 3 | | 4.54 | 29.03 | 9.68 | 8.43 | -3.04 | -2.50 | 2.80 | 2.14 | 2.55 | 3.89 |
| 4 | | 0.16 | -16.29 | -1.39 | -2.78 | -4.71 | -5.54 | -4.02 | 0.00 | -1.90 | -2.90 |
| 5 | | -1.16 | -34.38 | -5.36 | -6.92 | 11.63 | 9.57 | -1.78 | 0.04 | -0.73 | -0.62 |
| 6 | | 1.19 | -2.01 | 2.15 | 0.42 | 4.37 | 2.68 | 4.09 | 0.87 | 2.65 | 5.28 |
| 7 | | 5.79 | 40.28 | 13.21 | 11.28 | -7.51 | -10.02 | 2.31 | 0.33 | 1.51 | 3.53 |
| 8 | | 6.92 | 37.59 | 13.89 | 11.74 | 7.01 | 3.81 | -3.33 | 0.03 | -1.43 | -2.08 |
| 9 | | 1.37 | -7.55 | 1.95 | -0.44 | 11.60 | 9.68 | -2.89 | 0.36 | -1.02 | -1.60 |
| 10 | | 0.43 | -30.74 | -2.47 | -5.12 | -2.54 | -6.37 | 3.59 | 1.51 | 2.82 | 4.92 |
| 11 | | 0.63 | -5.17 | 2.20 | -0.74 | -2.33 | -3.72 | 4.76 | 1.74 | 3.55 | 6.13 |
| 12 | | 5.96 | 40.67 | 14.56 | 11.29 | 13.08 | 9.60 | -2.75 | -0.60 | -1.35 | -1.33 |
| 13 | | 6.64 | 44.63 | 16.08 | 12.46 | 3.32 | -0.70 | -5.26 | -1.11 | -2.83 | -3.78 |
| 14 | | 3.69 | 4.70 | 7.29 | 3.27 | -4.32 | -8.37 | 2.02 | 1.43 | 2.12 | 3.57 |
| 15 | | 0.95 | -24.93 | 0.42 | -4.03 | 11.17 | 5.75 | 4.85 | 1.31 | 3.51 | 6.47 |
| 16 | | 2.10 | -3.39 | 5.35 | 0.42 | 11.44 | 7.63 | 0.42 | 1.03 | 1.20 | 2.12 |
| 17 | | 7.09 | 44.24 | 17.86 | 12.40 | -3.60 | -8.97 | -5.34 | -0.92 | -2.61 | -3.54 |
| 18 | | 8.86 | 56.69 | 21.85 | 15.81 | 4.38 | -0.94 | -0.14 | 0.89 | 0.94 | 1.76 |
| 19 | | 6.20 | 21.07 | 14.18 | 7.51 | 17.62 | 10.74 | 5.27 | 1.82 | 4.17 | 7.29 |
| 20 | | 2.86 | -12.61 | 6.31 | -1.06 | 4.76 | -1.90 | 2.83 | 1.86 | 3.04 | 5.00 |
| 21 | | 4.52 | 4.48 | 11.11 | 2.96 | 0.29 | -8.13 | -4.76 | -0.74 | -1.98 | -2.44 |
| 22 | | 8.53 | 52.77 | 23.11 | 14.11 | 15.77 | 8.63 | -2.05 | 0.51 | 0.09 | 0.47 |
| 23 | | 11.87 | 74.62 | 30.19 | 20.25 | 14.15 | 5.00 | 4.37 | 1.64 | 3.96 | 7.10 |
| 24 | | 9.18 | 44.61 | 23.66 | 12.70 | 1.09 | -8.17 | 4.74 | 2.41 | 4.64 | 7.71 |
| 25 | | 4.52 | 7.82 | 14.41 | 2.31 | 11.56 | 1.91 | -2.76 | 0.27 | -0.05 | 0.49 |
| 26 | | 6.72 | 21.07 | 19.28 | 5.93 | 21.67 | 11.86 | -2.05 | 1.66 | 1.16 | 1.53 |
| 27 | | 9.75 | 69.98 | 30.84 | 16.12 | 5.24 | -4.34 | 3.49 | 1.97 | 4.26 | 7.45 |
| 28 | | 12.89 | 100.22 | 39.44 | 23.22 | 6.01 | -5.14 | 5.16 | 1.89 | 5.26 | 9.57 |
| 29 | | 11.14 | 78.88 | 35.44 | 17.57 | 22.96 | 10.82 | -0.43 | 1.15 | 2.33 | 4.49 |
| 30 | | 8.01 | 44.19 | 28.13 | 8.44 | 16.77 | 2.57 | -4.38 | -0.10 | 0.02 | 1.14 |
| 31 | | 7.43 | 51.73 | 30.06 | 8.38 | 5.31 | -7.86 | 1.85 | 1.79 | 4.41 | 8.08 |
| 32 | | 12.20 | 104.74 | 44.11 | 20.24 | 19.96 | 5.40 | 5.80 | 2.10 | 6.93 | 12.86 |
| 33 | | 17.51 | 147.20 | 56.92 | 30.64 | 26.92 | 10.68 | 2.37 | 2.21 | 5.74 | 10.42 |
| 34 | | 16.20 | 136.94 | 55.52 | 26.60 | 12.58 | -6.61 | -4.56 | -0.46 | 1.49 | 4.65 |
| 35 | | 11.28 | 104.59 | 47.50 | 15.69 | 17.01 | -0.91 | 0.15 | 1.45 | 5.47 | 10.76 |

| t | Варіант | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----------|---------|-------|--------|--------|-------|-------|-------|-------|------|-------|--------|
| | | | | | | | | | | | |
| 36 | | 11.88 | 113.21 | 50.99 | 16.00 | 32.68 | 13.98 | 7.29 | 3.61 | 10.92 | 19.57 |
| 37 | | 16.87 | 170.95 | 66.43 | 27.94 | 22.96 | 1.23 | 3.58 | 1.47 | 8.96 | 17.85 |
| 38 | | 21.94 | 226.08 | 81.57 | 39.26 | 17.65 | -5.61 | -2.40 | 0.54 | 6.66 | 14.29 |
| 39 | | 20.53 | 230.88 | 83.01 | 36.50 | 34.15 | 10.49 | -1.26 | 0.77 | 8.76 | 18.34 |
| 40 | | 17.22 | 209.73 | 78.63 | 27.51 | 35.89 | 11.91 | 6.90 | 3.25 | 15.79 | 30.03 |
| 41 | | 16.23 | 221.81 | 81.79 | 25.62 | 22.15 | -3.20 | 7.00 | 2.52 | 17.58 | 34.45 |
| 42 | | 20.95 | 288.42 | 98.87 | 37.16 | 32.79 | 6.96 | 1.45 | 1.72 | 16.98 | 34.20 |
| 43 | | 26.47 | 362.32 | 117.91 | 50.12 | 46.32 | 18.46 | -0.52 | 0.55 | 18.59 | 38.73 |
| 44 | | 25.62 | 390.17 | 124.01 | 49.57 | 33.55 | 6.87 | 7.14 | 2.33 | 27.24 | 54.41 |
| 45 | | 24.55 | 389.70 | 125.26 | 43.52 | 35.04 | 7.18 | 12.55 | 4.05 | 35.69 | 69.78 |
| 46 | | 23.94 | 416.25 | 131.42 | 41.68 | 53.93 | 26.24 | 8.39 | 2.78 | 39.06 | 78.00 |
| 47 | | 27.64 | 499.89 | 150.69 | 52.19 | 51.16 | 27.13 | 7.70 | 3.41 | 46.67 | 92.81 |
| 48 | | 34.94 | 603.87 | 176.34 | 68.23 | 42.70 | 19.29 | 14.80 | 3.97 | 60.09 | 119.36 |
| 49 | | 35.81 | 668.33 | 189.99 | 71.36 | 57.60 | 38.16 | 22.93 | 4.23 | 76.39 | 152.00 |
| 50 | | 34.46 | 700.72 | 196.37 | 66.22 | 69.64 | 57.59 | 26.42 | 5.68 | 94.21 | 186.51 |

Варіанти 11-20

| t | Варіант | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|-----------|---------|-------|-------|------|------|------|------|------|------|------|------|
| | | | | | | | | | | | |
| 1 | | -1.08 | -0.58 | 0.92 | 1.82 | 1.82 | 2.39 | 2.87 | 2.30 | 2.19 | 2.19 |
| 2 | | -0.39 | 0.00 | 1.80 | 2.61 | 2.78 | 2.31 | 2.68 | 2.31 | 2.17 | 2.19 |
| 3 | | 0.12 | 0.24 | 2.00 | 2.99 | 3.01 | 2.19 | 2.55 | 2.15 | 2.14 | 2.12 |
| 4 | | 0.72 | 0.51 | 2.58 | 3.37 | 3.00 | 2.09 | 2.51 | 2.12 | 2.13 | 2.05 |
| 5 | | 0.80 | 0.49 | 3.05 | 3.43 | 2.90 | 2.03 | 2.47 | 2.21 | 2.09 | 2.01 |
| 6 | | 0.87 | 0.79 | 3.10 | 3.78 | 3.49 | 2.04 | 2.43 | 2.14 | 2.03 | 2.02 |
| 7 | | 1.06 | 1.27 | 3.15 | 4.31 | 4.33 | 2.07 | 2.40 | 2.04 | 1.99 | 2.06 |
| 8 | | 1.04 | 1.29 | 3.46 | 4.35 | 4.35 | 2.05 | 2.39 | 2.12 | 1.96 | 2.05 |
| 9 | | 1.12 | 1.15 | 3.64 | 4.22 | 3.90 | 1.99 | 2.39 | 2.16 | 1.96 | 1.99 |
| 10 | | 1.44 | 1.29 | 3.58 | 4.37 | 3.87 | 1.96 | 2.40 | 2.06 | 1.98 | 1.96 |
| 11 | | 1.74 | 1.61 | 3.80 | 4.69 | 4.42 | 2.00 | 2.41 | 2.08 | 2.00 | 2.00 |
| 12 | | 1.63 | 1.61 | 3.85 | 4.69 | 4.76 | 2.03 | 2.40 | 2.16 | 2.02 | 2.03 |
| 13 | | 1.85 | 1.76 | 3.76 | 4.84 | 4.94 | 2.05 | 2.41 | 2.10 | 2.05 | 2.05 |
| 14 | | 2.07 | 1.75 | 3.76 | 4.83 | 4.63 | 2.00 | 2.42 | 2.05 | 2.06 | 2.00 |
| 15 | | 2.07 | 1.65 | 4.10 | 4.72 | 4.30 | 1.96 | 2.42 | 2.15 | 2.05 | 1.96 |

| t | Варіант | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|-----------|---------|------|------|------|------|------|------|------|------|------|------|
| | | | | | | | | | | | |
| 16 | | 2.00 | 1.80 | 4.15 | 4.86 | 4.62 | 1.96 | 2.39 | 2.13 | 2.00 | 1.96 |
| 17 | | 1.96 | 2.11 | 3.90 | 5.16 | 5.27 | 2.05 | 2.41 | 2.05 | 2.00 | 2.05 |
| 18 | | 1.67 | 1.97 | 3.84 | 5.02 | 5.22 | 2.06 | 2.41 | 2.10 | 1.98 | 2.06 |
| 19 | | 1.93 | 2.07 | 4.36 | 5.12 | 5.03 | 1.99 | 2.38 | 2.16 | 1.94 | 1.99 |
| 20 | | 1.75 | 1.68 | 3.87 | 4.72 | 4.37 | 1.95 | 2.39 | 2.08 | 1.96 | 1.95 |
| 21 | | 1.80 | 1.71 | 3.72 | 4.74 | 4.51 | 1.95 | 2.38 | 2.02 | 1.96 | 1.95 |
| 22 | | 1.95 | 1.97 | 4.07 | 5.00 | 5.12 | 2.03 | 2.41 | 2.16 | 2.02 | 2.03 |
| 23 | | 2.18 | 2.19 | 4.13 | 5.22 | 5.47 | 2.04 | 2.40 | 2.12 | 2.03 | 2.04 |
| 24 | | 2.11 | 1.91 | 3.73 | 4.93 | 4.93 | 2.00 | 2.41 | 2.03 | 2.05 | 2.00 |
| 25 | | 2.13 | 1.76 | 3.99 | 4.78 | 4.48 | 1.96 | 2.41 | 2.11 | 2.04 | 1.96 |
| 26 | | 2.02 | 1.78 | 4.14 | 4.80 | 4.56 | 1.96 | 2.39 | 2.16 | 2.01 | 1.96 |
| 27 | | 2.05 | 2.16 | 3.98 | 5.17 | 5.27 | 2.02 | 2.40 | 2.06 | 1.99 | 2.02 |
| 28 | | 1.66 | 1.99 | 3.69 | 5.01 | 5.29 | 2.05 | 2.39 | 2.06 | 1.96 | 2.05 |
| 29 | | 1.73 | 1.97 | 4.12 | 4.98 | 5.05 | 2.01 | 2.38 | 2.16 | 1.94 | 2.01 |
| 30 | | 1.81 | 1.81 | 4.02 | 4.82 | 4.56 | 1.97 | 2.40 | 2.12 | 1.96 | 1.97 |
| 31 | | 1.93 | 1.85 | 3.83 | 4.86 | 4.61 | 1.95 | 2.39 | 2.02 | 1.97 | 1.95 |
| 32 | | 1.90 | 1.91 | 3.95 | 4.92 | 4.99 | 2.02 | 2.41 | 2.13 | 2.01 | 2.02 |
| 33 | | 2.15 | 2.21 | 4.17 | 5.21 | 5.50 | 2.03 | 2.39 | 2.14 | 2.02 | 2.03 |
| 34 | | 2.16 | 2.04 | 3.80 | 5.04 | 5.16 | 2.01 | 2.40 | 2.04 | 2.04 | 2.01 |
| 35 | | 2.24 | 1.91 | 3.98 | 4.91 | 4.69 | 1.96 | 2.41 | 2.07 | 2.04 | 1.96 |
| 36 | | 2.11 | 1.82 | 4.21 | 4.83 | 4.55 | 1.96 | 2.41 | 2.18 | 2.03 | 1.96 |
| 37 | | 2.15 | 2.18 | 4.13 | 5.18 | 5.21 | 2.02 | 2.42 | 2.11 | 2.01 | 2.02 |
| 38 | | 2.10 | 2.41 | 4.04 | 5.41 | 5.70 | 2.05 | 2.39 | 2.04 | 1.97 | 2.05 |
| 39 | | 1.91 | 2.21 | 4.23 | 5.21 | 5.38 | 2.03 | 2.39 | 2.15 | 1.96 | 2.03 |
| 40 | | 2.05 | 2.12 | 4.36 | 5.12 | 4.94 | 1.96 | 2.38 | 2.13 | 1.95 | 1.96 |
| 41 | | 2.12 | 2.04 | 4.05 | 5.04 | 4.75 | 1.94 | 2.38 | 2.02 | 1.96 | 1.94 |
| 42 | | 1.89 | 1.88 | 3.87 | 4.88 | 4.86 | 2.01 | 2.41 | 2.11 | 2.01 | 2.01 |
| 43 | | 2.17 | 2.24 | 4.26 | 5.24 | 5.52 | 2.06 | 2.42 | 2.19 | 2.04 | 2.06 |
| 44 | | 1.89 | 1.84 | 3.62 | 4.84 | 5.05 | 2.03 | 2.41 | 2.07 | 2.05 | 2.03 |
| 45 | | 2.18 | 1.90 | 3.83 | 4.90 | 4.76 | 1.99 | 2.42 | 2.06 | 2.06 | 1.99 |
| 46 | | 2.08 | 1.75 | 4.12 | 4.75 | 4.45 | 1.93 | 2.38 | 2.14 | 2.01 | 1.93 |
| 47 | | 2.03 | 1.96 | 4.08 | 4.97 | 4.90 | 1.96 | 2.38 | 2.10 | 1.98 | 1.96 |
| 48 | | 1.86 | 2.13 | 3.76 | 5.13 | 5.38 | 2.06 | 2.41 | 2.04 | 1.99 | 2.06 |
| 49 | | 2.14 | 2.47 | 4.36 | 5.47 | 5.71 | 2.03 | 2.39 | 2.11 | 1.95 | 2.03 |
| 50 | | 1.83 | 1.97 | 4.21 | 4.97 | 4.87 | 1.98 | 2.38 | 2.16 | 1.94 | 1.98 |

ДОДАТОК Б

Показники динаміки розвитку економічних процесів

Спочатку у часовому ряді необхідно виявити закономірності динаміки досліджуваних явищ. Закономірності можуть бути виявлені за допомогою аналітичних показників, тобто аналізу абсолютної швидкості та інтенсивності розвитку явищ.

Для дослідження часового ряду необхідно, щоб рівні ряду були однорідні. Причиною неоднорідності даних можуть бути дані різних регіонів, об'єднань, різних років, різних масштабів, різних структур сукупності та інші.

При дослідженні часових рядів розраховують абсолютний приріст:

$$\Delta y_i = y_i - y_{i-k}, \quad i = 2, 3, \dots, n, \quad k = 1, 2, \dots, n - 1$$

— визначає початковий рівень і може бути різним. Якщо $k = 1$, то отримаємо ланцюгові показники, а при $k = i - 1$ маємо базисні показники з початковим рівнем ряду в якості базисного.

Отже, абсолютний приріст є величина зміни показника за час між порівнювальними періодами.

Середнім абсолютним приростом, тобто швидкістю зміни абсолютного приросту, називають величину

$$\overline{\Delta y}_i = \frac{y_i - y_{i-k}}{k}.$$

Якщо $i = n$, $k = n - 1$, то отримаємо середній абсолютний приріст за весь період спостережень

$$\overline{\Delta y} = \frac{y_n - y_1}{n - 1},$$

що характеризує середню швидкість зміни часового ряду.

Для визначення відносної швидкості зміни явища за одиницю часу використовують відносні показники.

Коефіцієнт зростання для i -го періоду розраховується за формулою

$$K_{i(ЗР)} = \frac{y_i}{y_{i-k}}.$$

Якщо $K_{i(ЗР)} > 1$, то рівень зростає, при $K_{i(ЗР)} < 1$ — зменшується, а при $K_{i(ЗР)} = 1$ рівень не змінюється.

Коефіцієнт приросту розраховується так:

$$K_{i(ПР)} = K_{i(ЗР)} - 1 = \frac{y_i - y_{i-k}}{y_{i-k}}.$$

Також використовують показники темпу зростання і темпу приросту:

$$T_{i(ЗР)} = \frac{y_i}{y_{i-k}} 100\%, \quad T_{i(ПР)} = T_{i(ЗР)} - 100\% = \frac{y_i - y_{i-k}}{y_{i-k}} 100\%.$$

Темп зростання показує, на скільки відсотків рівень одного періоду змінився у порівнянні з рівнем іншого періоду.

Може статися так, що зменшення темпу приросту не супроводжується зменшенням абсолютних приростів.

Середній темп зростання характеризує середню швидкість зміни явища за весь період:

$$\bar{T}_{(ЗР)} = \sqrt[n-1]{\frac{y_n}{y_1}} 100\%.$$

Якщо ряд має сильні коливання, то ця формула дає не дуже точні результати, і тому краще користуватися іншою:

$$\bar{T}_{(ЗР)} = \sqrt[n-1]{\frac{\hat{y}_n}{\hat{y}_1}} 100\%,$$

де \hat{y}_n і \hat{y}_1 — згладжені за рівнянням тренду початковий і кінцевий рівні часового ряду.

Ще використовується такий показник, як середній рівень ряду: $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$.

ДОДАТОК В

Таблиці математичної статистики

В.1 q -квантилі стандартного нормального розподілу

Табл. В.1. Значення q -квантилей u_q стандартного нормального розподілу

| q | u_q | q | u_q | q | u_q | q | u_q |
|-------------|----------|-------------|----------|--------------|----------|--------------|----------|
| 0.50 | 0.000000 | 0.70 | 0.524401 | 0.90 | 1.281552 | 0.983 | 2.120072 |
| 0.51 | 0.025069 | 0.71 | 0.553385 | 0.91 | 1.340755 | 0.984 | 2.144411 |
| 0.52 | 0.050154 | 0.72 | 0.582842 | 0.92 | 1.405072 | 0.985 | 2.170090 |
| 0.53 | 0.075270 | 0.73 | 0.612813 | 0.93 | 1.475791 | 0.986 | 2.197286 |
| 0.54 | 0.100434 | 0.74 | 0.643345 | 0.94 | 1.554774 | 0.987 | 2.226212 |
| 0.55 | 0.125661 | 0.75 | 0.674490 | 0.95 | 1.644854 | 0.988 | 2.257129 |
| 0.56 | 0.150969 | 0.76 | 0.706303 | 0.96 | 1.750686 | 0.989 | 2.290368 |
| 0.57 | 0.176374 | 0.77 | 0.738847 | 0.97 | 1.880794 | 0.990 | 2.326348 |
| 0.58 | 0.201893 | 0.78 | 0.772193 | 0.971 | 1.895698 | 0.991 | 2.365618 |
| 0.59 | 0.227545 | 0.79 | 0.806421 | 0.972 | 1.911036 | 0.992 | 2.408916 |
| 0.60 | 0.253347 | 0.80 | 0.841621 | 0.973 | 1.926837 | 0.993 | 2.457263 |
| 0.61 | 0.279319 | 0.81 | 0.877896 | 0.974 | 1.943134 | 0.994 | 2.512144 |
| 0.62 | 0.305481 | 0.82 | 0.915365 | 0.975 | 1.959964 | 0.995 | 2.575829 |
| 0.63 | 0.331853 | 0.83 | 0.954165 | 0.976 | 1.977368 | 0.996 | 2.652070 |
| 0.64 | 0.358459 | 0.84 | 0.994458 | 0.977 | 1.995393 | 0.997 | 2.747781 |
| 0.65 | 0.385320 | 0.85 | 1.036433 | 0.978 | 2.014091 | 0.998 | 2.878162 |
| 0.66 | 0.412463 | 0.86 | 1.080319 | 0.979 | 2.033520 | 0.999 | 3.090232 |
| 0.67 | 0.439913 | 0.87 | 1.126391 | 0.980 | 2.053749 | | |
| 0.68 | 0.467699 | 0.88 | 1.174987 | 0.981 | 2.074855 | | |
| 0.69 | 0.495850 | 0.89 | 1.226528 | 0.982 | 2.096927 | | |

Примітки.

1. При знаходженні q -квантилей для значень $q < 0.5$ слід скористатися співвідношенням $u_q = -u_{1-q}$. Наприклад, $u_{0.4} = -u_{1-0.4} = -u_{0.6} = -0.253347$.
2. Значення q -квантилей u_q можна розрахувати у Microsoft Excel за допомогою функції НОРМ.ОБР(q ; 0; 1) категорії «Статистичні». Наприклад, щоб розрахувати $u_{0.95}$, слід задати формулу '=НОРМ.ОБР(0.95; 0; 1)'.

В.2 Розподіл Пірсона χ^2

Табл. В.2. Критичні значення χ^2 -розподілу Пірсона $\chi^2(\alpha; \nu)$ для рівнів значущості α та ν ступенів вільності

| $\alpha \backslash \nu$ | 0.995 | 0.99 | 0.975 | 0.95 | 0.9 | 0.1 | 0.05 | 0.025 | 0.01 | 0.005 |
|-------------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 1 | 0.00004 | 0.00016 | 0.00098 | 0.00393 | 0.01579 | 2.70554 | 3.84146 | 5.02389 | 6.63490 | 7.87944 |
| 2 | 0.01003 | 0.02010 | 0.05064 | 0.10259 | 0.21072 | 4.60517 | 5.99146 | 7.37776 | 9.21034 | 10.59663 |
| 3 | 0.07172 | 0.11483 | 0.21580 | 0.35185 | 0.58437 | 6.25139 | 7.81473 | 9.34840 | 11.34487 | 12.83816 |
| 4 | 0.20699 | 0.29711 | 0.48442 | 0.71072 | 1.06362 | 7.77944 | 9.48773 | 11.14329 | 13.27670 | 14.86026 |
| 5 | 0.41174 | 0.55430 | 0.83121 | 1.14548 | 1.61031 | 9.23636 | 11.07050 | 12.83250 | 15.08627 | 16.74960 |
| 6 | 0.67573 | 0.87209 | 1.23734 | 1.63538 | 2.20413 | 10.64464 | 12.59159 | 14.44938 | 16.81189 | 18.54758 |
| 7 | 0.98926 | 1.23904 | 1.68987 | 2.16735 | 2.83311 | 12.01704 | 14.06714 | 16.01276 | 18.47531 | 20.27774 |
| 8 | 1.34441 | 1.64650 | 2.17973 | 2.73264 | 3.48954 | 13.36157 | 15.50731 | 17.53455 | 20.09024 | 21.95495 |
| 9 | 1.73493 | 2.08790 | 2.70039 | 3.32511 | 4.16816 | 14.68366 | 16.91898 | 19.02277 | 21.66599 | 23.58935 |
| 10 | 2.15586 | 2.55821 | 3.24697 | 3.94030 | 4.86518 | 15.98718 | 18.30704 | 20.48318 | 23.20925 | 25.18818 |
| 11 | 2.60322 | 3.05348 | 3.81575 | 4.57481 | 5.57778 | 17.27501 | 19.67514 | 21.92005 | 24.72497 | 26.75685 |
| 12 | 3.07382 | 3.57057 | 4.40379 | 5.22603 | 6.30380 | 18.54935 | 21.02607 | 23.33666 | 26.21697 | 28.29952 |
| 13 | 3.56503 | 4.10692 | 5.00875 | 5.89186 | 7.04150 | 19.81193 | 22.36203 | 24.73560 | 27.68825 | 29.81947 |
| 14 | 4.07467 | 4.66043 | 5.62873 | 6.57063 | 7.78953 | 21.06414 | 23.68479 | 26.11895 | 29.14124 | 31.31935 |
| 15 | 4.60092 | 5.22935 | 6.26214 | 7.26094 | 8.54676 | 22.30713 | 24.99579 | 27.48839 | 30.57791 | 32.80132 |
| 16 | 5.14221 | 5.81221 | 6.90766 | 7.96165 | 9.31224 | 23.54183 | 26.29623 | 28.84535 | 31.99993 | 34.26719 |
| 17 | 5.69722 | 6.40776 | 7.56419 | 8.67176 | 10.08519 | 24.76904 | 27.58711 | 30.19101 | 33.40866 | 35.71847 |
| 18 | 6.26480 | 7.01491 | 8.23075 | 9.39046 | 10.86494 | 25.98942 | 28.86930 | 31.52638 | 34.80531 | 37.15645 |
| 19 | 6.84397 | 7.63273 | 8.90652 | 10.11701 | 11.65091 | 27.20357 | 30.14353 | 32.85233 | 36.19087 | 38.58226 |
| 20 | 7.43384 | 8.26040 | 9.59078 | 10.85081 | 12.44261 | 28.41198 | 31.41043 | 34.16961 | 37.56623 | 39.99685 |
| 21 | 8.03365 | 8.89720 | 10.28290 | 11.59131 | 13.23960 | 29.61509 | 32.67057 | 35.47888 | 38.93217 | 41.40106 |
| 22 | 8.64272 | 9.54249 | 10.98232 | 12.33801 | 14.04149 | 30.81328 | 33.92444 | 36.78071 | 40.28936 | 42.79565 |
| 23 | 9.26042 | 10.19572 | 11.68855 | 13.09051 | 14.84796 | 32.00690 | 35.17246 | 38.07563 | 41.63840 | 44.18128 |
| 24 | 9.88623 | 10.85636 | 12.40115 | 13.84843 | 15.65868 | 33.19624 | 36.41503 | 39.36408 | 42.97982 | 45.55851 |
| 25 | 10.51965 | 11.52398 | 13.11972 | 14.61141 | 16.47341 | 34.38159 | 37.65248 | 40.64647 | 44.31410 | 46.92789 |
| 26 | 11.16024 | 12.19815 | 13.84390 | 15.37916 | 17.29188 | 35.56317 | 38.88514 | 41.92317 | 45.64168 | 48.28988 |
| 27 | 11.80759 | 12.87850 | 14.57338 | 16.15140 | 18.11390 | 36.74122 | 40.11327 | 43.19451 | 46.96294 | 49.64492 |
| 28 | 12.46134 | 13.56471 | 15.30786 | 16.92788 | 18.93924 | 37.91592 | 41.33714 | 44.46079 | 48.27824 | 50.99338 |
| 29 | 13.12115 | 14.25645 | 16.04707 | 17.70837 | 19.76774 | 39.08747 | 42.55697 | 45.72229 | 49.58788 | 52.33562 |
| 30 | 13.78672 | 14.95346 | 16.79077 | 18.49266 | 20.59923 | 40.25602 | 43.77297 | 46.97924 | 50.89218 | 53.67196 |
| 40 | 20.70654 | 22.16426 | 24.43304 | 26.50930 | 29.05052 | 51.80506 | 55.75848 | 59.34171 | 63.69074 | 66.76596 |
| 50 | 27.99075 | 29.70668 | 32.35736 | 34.76425 | 37.68865 | 63.16712 | 67.50481 | 71.42020 | 76.15389 | 79.48998 |
| 60 | 35.53449 | 37.48485 | 40.48175 | 43.18796 | 46.45889 | 74.39701 | 79.08194 | 83.29767 | 88.37942 | 91.95170 |
| 70 | 43.27518 | 45.44172 | 48.75756 | 51.73928 | 55.32894 | 85.52704 | 90.53123 | 95.02318 | 100.4252 | 104.2149 |
| 80 | 51.17193 | 53.54008 | 57.15317 | 60.39148 | 64.27784 | 96.57820 | 101.8795 | 106.6286 | 112.3288 | 116.3211 |
| 90 | 59.19630 | 61.75408 | 65.64662 | 69.12603 | 73.29109 | 107.5650 | 113.1453 | 118.1359 | 124.1163 | 128.2989 |
| 100 | 67.32756 | 70.06489 | 74.22193 | 77.92947 | 82.35814 | 118.4980 | 124.3421 | 129.5612 | 135.8067 | 140.1695 |

Примітка. Критичні значення $\chi^2(\alpha; \nu)$ можна розрахувати у Microsoft Excel за допомогою функції `ХИ2.ОБР(1 - α ; ν)` категорії «Статистичні». Наприклад, для розрахунку $\chi^2(0.1; 30)$ слід задати формулу `'=ХИ2.ОБР(0.9; 30)'`.

В.3 F -розподіл Фішера

Табл. В.3. Критичні значення F -розподілу Фішера $F(\alpha; \nu_1; \nu_2)$ для рівнів значущості α із числом ступенів вільності чисельника ν_1 та знаменника ν_2

| $\nu_1 \backslash \nu_2$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 15 | 20 | 25 | 30 | 40 | 60 | 120 |
|--------------------------|-----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | $\alpha = 0.1$ | | | | | | | | | | | | | | | | | | |
| 1 | 39.86 | 49.50 | 53.59 | 55.83 | 57.24 | 58.20 | 58.91 | 59.44 | 59.86 | 60.19 | 60.47 | 60.71 | 61.22 | 61.74 | 62.05 | 62.26 | 62.53 | 62.79 | 63.06 |
| 2 | 8.53 | 9.00 | 9.16 | 9.24 | 9.29 | 9.33 | 9.35 | 9.37 | 9.38 | 9.39 | 9.40 | 9.41 | 9.42 | 9.44 | 9.45 | 9.46 | 9.47 | 9.47 | 9.48 |
| 3 | 5.54 | 5.46 | 5.39 | 5.34 | 5.31 | 5.28 | 5.27 | 5.25 | 5.24 | 5.23 | 5.22 | 5.22 | 5.20 | 5.18 | 5.17 | 5.17 | 5.16 | 5.15 | 5.14 |
| 4 | 4.54 | 4.32 | 4.19 | 4.11 | 4.05 | 4.01 | 3.98 | 3.95 | 3.94 | 3.92 | 3.91 | 3.90 | 3.87 | 3.84 | 3.83 | 3.82 | 3.80 | 3.79 | 3.78 |
| 5 | 4.06 | 3.78 | 3.62 | 3.52 | 3.45 | 3.40 | 3.37 | 3.34 | 3.32 | 3.30 | 3.28 | 3.27 | 3.24 | 3.21 | 3.19 | 3.17 | 3.16 | 3.14 | 3.12 |
| 6 | 3.78 | 3.46 | 3.29 | 3.18 | 3.11 | 3.05 | 3.01 | 2.98 | 2.96 | 2.94 | 2.92 | 2.90 | 2.87 | 2.84 | 2.81 | 2.80 | 2.78 | 2.76 | 2.74 |
| 7 | 3.59 | 3.26 | 3.07 | 2.96 | 2.88 | 2.83 | 2.78 | 2.75 | 2.72 | 2.70 | 2.68 | 2.67 | 2.63 | 2.59 | 2.57 | 2.56 | 2.54 | 2.51 | 2.49 |
| 8 | 3.46 | 3.11 | 2.92 | 2.81 | 2.73 | 2.67 | 2.62 | 2.59 | 2.56 | 2.54 | 2.52 | 2.50 | 2.46 | 2.42 | 2.40 | 2.38 | 2.36 | 2.34 | 2.32 |
| 9 | 3.36 | 3.01 | 2.81 | 2.69 | 2.61 | 2.55 | 2.51 | 2.47 | 2.44 | 2.42 | 2.40 | 2.38 | 2.34 | 2.30 | 2.27 | 2.25 | 2.23 | 2.21 | 2.18 |
| 10 | 3.29 | 2.92 | 2.73 | 2.61 | 2.52 | 2.46 | 2.41 | 2.38 | 2.35 | 2.32 | 2.30 | 2.28 | 2.24 | 2.20 | 2.17 | 2.16 | 2.13 | 2.11 | 2.08 |
| 11 | 3.23 | 2.86 | 2.66 | 2.54 | 2.45 | 2.39 | 2.34 | 2.30 | 2.27 | 2.25 | 2.23 | 2.21 | 2.17 | 2.12 | 2.10 | 2.08 | 2.05 | 2.03 | 2.00 |
| 12 | 3.18 | 2.81 | 2.61 | 2.48 | 2.39 | 2.33 | 2.28 | 2.24 | 2.21 | 2.19 | 2.17 | 2.15 | 2.10 | 2.06 | 2.03 | 2.01 | 1.99 | 1.96 | 1.93 |
| 13 | 3.14 | 2.76 | 2.56 | 2.43 | 2.35 | 2.28 | 2.23 | 2.20 | 2.16 | 2.14 | 2.12 | 2.10 | 2.05 | 2.01 | 1.98 | 1.96 | 1.93 | 1.90 | 1.88 |
| 14 | 3.10 | 2.73 | 2.52 | 2.39 | 2.31 | 2.24 | 2.19 | 2.15 | 2.12 | 2.10 | 2.07 | 2.05 | 2.01 | 1.96 | 1.93 | 1.91 | 1.89 | 1.86 | 1.83 |
| 15 | 3.07 | 2.70 | 2.49 | 2.36 | 2.27 | 2.21 | 2.16 | 2.12 | 2.09 | 2.06 | 2.04 | 2.02 | 1.97 | 1.92 | 1.89 | 1.87 | 1.85 | 1.82 | 1.79 |
| 16 | 3.05 | 2.67 | 2.46 | 2.33 | 2.24 | 2.18 | 2.13 | 2.09 | 2.06 | 2.03 | 2.01 | 1.99 | 1.94 | 1.89 | 1.86 | 1.84 | 1.81 | 1.78 | 1.75 |
| 17 | 3.03 | 2.64 | 2.44 | 2.31 | 2.22 | 2.15 | 2.10 | 2.06 | 2.03 | 2.00 | 1.98 | 1.96 | 1.91 | 1.86 | 1.83 | 1.81 | 1.78 | 1.75 | 1.72 |
| 18 | 3.01 | 2.62 | 2.42 | 2.29 | 2.20 | 2.13 | 2.08 | 2.04 | 2.00 | 1.98 | 1.95 | 1.93 | 1.89 | 1.84 | 1.80 | 1.78 | 1.75 | 1.72 | 1.69 |
| 19 | 2.99 | 2.61 | 2.40 | 2.27 | 2.18 | 2.11 | 2.06 | 2.02 | 1.98 | 1.96 | 1.93 | 1.91 | 1.86 | 1.81 | 1.78 | 1.76 | 1.73 | 1.70 | 1.67 |
| 20 | 2.97 | 2.59 | 2.38 | 2.25 | 2.16 | 2.09 | 2.04 | 2.00 | 1.96 | 1.94 | 1.91 | 1.89 | 1.84 | 1.79 | 1.76 | 1.74 | 1.71 | 1.68 | 1.64 |
| 21 | 2.96 | 2.57 | 2.36 | 2.23 | 2.14 | 2.08 | 2.02 | 1.98 | 1.95 | 1.92 | 1.90 | 1.87 | 1.83 | 1.78 | 1.74 | 1.72 | 1.69 | 1.66 | 1.62 |
| 22 | 2.95 | 2.56 | 2.35 | 2.22 | 2.13 | 2.06 | 2.01 | 1.97 | 1.93 | 1.90 | 1.88 | 1.86 | 1.81 | 1.76 | 1.73 | 1.70 | 1.67 | 1.64 | 1.60 |
| 23 | 2.94 | 2.55 | 2.34 | 2.21 | 2.11 | 2.05 | 1.99 | 1.95 | 1.92 | 1.89 | 1.87 | 1.84 | 1.80 | 1.74 | 1.71 | 1.69 | 1.66 | 1.62 | 1.59 |
| 24 | 2.93 | 2.54 | 2.33 | 2.19 | 2.10 | 2.04 | 1.98 | 1.94 | 1.91 | 1.88 | 1.85 | 1.83 | 1.78 | 1.73 | 1.70 | 1.67 | 1.64 | 1.61 | 1.57 |
| 25 | 2.92 | 2.53 | 2.32 | 2.18 | 2.09 | 2.02 | 1.97 | 1.93 | 1.89 | 1.87 | 1.84 | 1.82 | 1.77 | 1.72 | 1.68 | 1.66 | 1.63 | 1.59 | 1.56 |
| 26 | 2.91 | 2.52 | 2.31 | 2.17 | 2.08 | 2.01 | 1.96 | 1.92 | 1.88 | 1.86 | 1.83 | 1.81 | 1.76 | 1.71 | 1.67 | 1.65 | 1.61 | 1.58 | 1.54 |
| 27 | 2.90 | 2.51 | 2.30 | 2.17 | 2.07 | 2.00 | 1.95 | 1.91 | 1.87 | 1.85 | 1.82 | 1.80 | 1.75 | 1.70 | 1.66 | 1.64 | 1.60 | 1.57 | 1.53 |
| 28 | 2.89 | 2.50 | 2.29 | 2.16 | 2.06 | 2.00 | 1.94 | 1.90 | 1.87 | 1.84 | 1.81 | 1.79 | 1.74 | 1.69 | 1.65 | 1.63 | 1.59 | 1.56 | 1.52 |
| 29 | 2.89 | 2.50 | 2.28 | 2.15 | 2.06 | 1.99 | 1.93 | 1.89 | 1.86 | 1.83 | 1.80 | 1.78 | 1.73 | 1.68 | 1.64 | 1.62 | 1.58 | 1.55 | 1.51 |
| 30 | 2.88 | 2.49 | 2.28 | 2.14 | 2.05 | 1.98 | 1.93 | 1.88 | 1.85 | 1.82 | 1.79 | 1.77 | 1.72 | 1.67 | 1.63 | 1.61 | 1.57 | 1.54 | 1.50 |
| 40 | 2.84 | 2.44 | 2.23 | 2.09 | 2.00 | 1.93 | 1.87 | 1.83 | 1.79 | 1.76 | 1.74 | 1.71 | 1.66 | 1.61 | 1.57 | 1.54 | 1.51 | 1.47 | 1.42 |
| 60 | 2.79 | 2.39 | 2.18 | 2.04 | 1.95 | 1.87 | 1.82 | 1.77 | 1.74 | 1.71 | 1.68 | 1.66 | 1.60 | 1.54 | 1.50 | 1.48 | 1.44 | 1.40 | 1.35 |
| 120 | 2.75 | 2.35 | 2.13 | 1.99 | 1.90 | 1.82 | 1.77 | 1.72 | 1.68 | 1.65 | 1.63 | 1.60 | 1.55 | 1.48 | 1.44 | 1.41 | 1.37 | 1.32 | 1.26 |
| | $\alpha = 0.05$ | | | | | | | | | | | | | | | | | | |
| 2 | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 19.35 | 19.37 | 19.38 | 19.40 | 19.40 | 19.41 | 19.43 | 19.45 | 19.46 | 19.46 | 19.47 | 19.48 | 19.49 |
| 3 | 10.13 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.89 | 8.85 | 8.81 | 8.79 | 8.76 | 8.74 | 8.70 | 8.66 | 8.63 | 8.62 | 8.59 | 8.57 | 8.55 |
| 4 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.09 | 6.04 | 6.00 | 5.96 | 5.94 | 5.91 | 5.86 | 5.80 | 5.77 | 5.75 | 5.72 | 5.69 | 5.66 |
| 5 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 | 4.74 | 4.70 | 4.68 | 4.62 | 4.56 | 4.52 | 4.50 | 4.46 | 4.43 | 4.40 |
| 6 | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.21 | 4.15 | 4.10 | 4.06 | 4.03 | 4.00 | 3.94 | 3.87 | 3.83 | 3.81 | 3.77 | 3.74 | 3.70 |
| 7 | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.79 | 3.73 | 3.68 | 3.64 | 3.60 | 3.57 | 3.51 | 3.44 | 3.40 | 3.38 | 3.34 | 3.30 | 3.27 |
| 8 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.50 | 3.44 | 3.39 | 3.35 | 3.31 | 3.28 | 3.22 | 3.15 | 3.11 | 3.08 | 3.04 | 3.01 | 2.97 |
| 9 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.29 | 3.23 | 3.18 | 3.14 | 3.10 | 3.07 | 3.01 | 2.94 | 2.89 | 2.86 | 2.83 | 2.79 | 2.75 |
| 10 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 | 2.98 | 2.94 | 2.91 | 2.85 | 2.77 | 2.73 | 2.70 | 2.66 | 2.62 | 2.58 |
| 11 | 4.84 | 3.98 | 3.59 | 3.36 | 3.20 | 3.09 | 3.01 | 2.95 | 2.90 | 2.85 | 2.82 | 2.79 | 2.72 | 2.65 | 2.60 | 2.57 | 2.53 | 2.49 | 2.45 |
| 12 | 4.75 | 3.89 | 3.49 | 3.26 | 3.11 | 3.00 | 2.91 | 2.85 | 2.80 | 2.75 | 2.72 | 2.69 | 2.62 | 2.54 | 2.50 | 2.47 | 2.43 | 2.38 | 2.34 |
| 13 | 4.67 | 3.81 | 3.41 | 3.18 | 3.03 | 2.92 | 2.83 | 2.77 | 2.71 | 2.67 | 2.63 | 2.60 | 2.53 | 2.46 | 2.41 | 2.38 | 2.34 | 2.30 | 2.25 |
| 14 | 4.60 | 3.74 | 3.34 | 3.11 | 2.96 | 2.85 | 2.76 | 2.70 | 2.65 | 2.60 | 2.57 | 2.53 | 2.46 | 2.39 | 2.34 | 2.31 | 2.27 | 2.22 | 2.18 |
| 15 | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.71 | 2.64 | 2.59 | 2.54 | 2.51 | 2.48 | 2.40 | 2.33 | 2.28 | 2.25 | 2.20 | 2.16 | 2.11 |
| 16 | 4.49 | 3.63 | 3.24 | 3.01 | 2.85 | 2.74 | 2.66 | 2.59 | 2.54 | 2.49 | 2.46 | 2.42 | 2.35 | 2.28 | 2.23 | 2.19 | 2.15 | 2.11 | 2.06 |
| 17 | 4.45 | 3.59 | 3.20 | 2.96 | 2.81 | 2.70 | 2.61 | 2.55 | 2.49 | 2.45 | 2.41 | 2.38 | 2.31 | 2.23 | 2.18 | 2.15 | 2.10 | 2.06 | 2.01 |

Табл. В.3. Критичні значення F -розподілу Фішера $F(\alpha; \nu_1; \nu_2)$ для рівнів значущості α із числом ступенів вільності чисельника ν_1 та знаменника ν_2

| $\nu_1 \backslash \nu_2$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 15 | 20 | 25 | 30 | 40 | 60 | 120 |
|--------------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 18 | 4.41 | 3.55 | 3.16 | 2.93 | 2.77 | 2.66 | 2.58 | 2.51 | 2.46 | 2.41 | 2.37 | 2.34 | 2.27 | 2.19 | 2.14 | 2.11 | 2.06 | 2.02 | 1.97 |
| 19 | 4.38 | 3.52 | 3.13 | 2.90 | 2.74 | 2.63 | 2.54 | 2.48 | 2.42 | 2.38 | 2.34 | 2.31 | 2.23 | 2.16 | 2.11 | 2.07 | 2.03 | 1.98 | 1.93 |
| 20 | 4.35 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.51 | 2.45 | 2.39 | 2.35 | 2.31 | 2.28 | 2.20 | 2.12 | 2.07 | 2.04 | 1.99 | 1.95 | 1.90 |
| 21 | 4.32 | 3.47 | 3.07 | 2.84 | 2.68 | 2.57 | 2.49 | 2.42 | 2.37 | 2.32 | 2.28 | 2.25 | 2.18 | 2.10 | 2.05 | 2.01 | 1.96 | 1.92 | 1.87 |
| 22 | 4.30 | 3.44 | 3.05 | 2.82 | 2.66 | 2.55 | 2.46 | 2.40 | 2.34 | 2.30 | 2.26 | 2.23 | 2.15 | 2.07 | 2.02 | 1.98 | 1.94 | 1.89 | 1.84 |
| 23 | 4.28 | 3.42 | 3.03 | 2.80 | 2.64 | 2.53 | 2.44 | 2.37 | 2.32 | 2.27 | 2.24 | 2.20 | 2.13 | 2.05 | 2.00 | 1.96 | 1.91 | 1.86 | 1.81 |
| 24 | 4.26 | 3.40 | 3.01 | 2.78 | 2.62 | 2.51 | 2.42 | 2.36 | 2.30 | 2.25 | 2.22 | 2.18 | 2.11 | 2.03 | 1.97 | 1.94 | 1.89 | 1.84 | 1.79 |
| 25 | 4.24 | 3.39 | 2.99 | 2.76 | 2.60 | 2.49 | 2.40 | 2.34 | 2.28 | 2.24 | 2.20 | 2.16 | 2.09 | 2.01 | 1.96 | 1.92 | 1.87 | 1.82 | 1.77 |
| 26 | 4.23 | 3.37 | 2.98 | 2.74 | 2.59 | 2.47 | 2.39 | 2.32 | 2.27 | 2.22 | 2.18 | 2.15 | 2.07 | 1.99 | 1.94 | 1.90 | 1.85 | 1.80 | 1.75 |
| 27 | 4.21 | 3.35 | 2.96 | 2.73 | 2.57 | 2.46 | 2.37 | 2.31 | 2.25 | 2.20 | 2.17 | 2.13 | 2.06 | 1.97 | 1.92 | 1.88 | 1.84 | 1.79 | 1.73 |
| 28 | 4.20 | 3.34 | 2.95 | 2.71 | 2.56 | 2.45 | 2.36 | 2.29 | 2.24 | 2.19 | 2.15 | 2.12 | 2.04 | 1.96 | 1.91 | 1.87 | 1.82 | 1.77 | 1.71 |
| 29 | 4.18 | 3.33 | 2.93 | 2.70 | 2.55 | 2.43 | 2.35 | 2.28 | 2.22 | 2.18 | 2.14 | 2.10 | 2.03 | 1.94 | 1.89 | 1.85 | 1.81 | 1.75 | 1.70 |
| 30 | 4.17 | 3.32 | 2.92 | 2.69 | 2.53 | 2.42 | 2.33 | 2.27 | 2.21 | 2.16 | 2.13 | 2.09 | 2.01 | 1.93 | 1.88 | 1.84 | 1.79 | 1.74 | 1.68 |
| 40 | 4.08 | 3.23 | 2.84 | 2.61 | 2.45 | 2.34 | 2.25 | 2.18 | 2.12 | 2.08 | 2.04 | 2.00 | 1.92 | 1.84 | 1.78 | 1.74 | 1.69 | 1.64 | 1.58 |
| 60 | 4.00 | 3.15 | 2.76 | 2.53 | 2.37 | 2.25 | 2.17 | 2.10 | 2.04 | 1.99 | 1.95 | 1.92 | 1.84 | 1.75 | 1.69 | 1.65 | 1.59 | 1.53 | 1.47 |
| 120 | 3.92 | 3.07 | 2.68 | 2.45 | 2.29 | 2.18 | 2.09 | 2.02 | 1.96 | 1.91 | 1.87 | 1.83 | 1.75 | 1.66 | 1.60 | 1.55 | 1.50 | 1.43 | 1.35 |

Примітки.

1. При обчисленні критичних значень для $\alpha \geq 0.9$ слід скористатися тотожністю

$$F(\alpha; \nu_1; \nu_2) = (F(1 - \alpha; \nu_1; \nu_2))^{-1}.$$

2. Критичні значення $F(\alpha; \nu_1; \nu_2)$ можна розрахувати у Microsoft Excel за допомогою функції $\text{F.ОБР}(1 - \alpha; \nu_1; \nu_2)$ категорії «Статистичні». Наприклад, для розрахунку $F(0.1; 5; 10)$ слід задати формулу $'=\text{F.ОБР}(0.9; 5; 10)'$.

В.4 Розподіл Ст'юдента

Табл. В.4. Критичні значення $t(\alpha; \nu)$ розподілу Ст'юдента для рівнів значущості α з ν ступенями вільності

| $\nu \backslash \alpha$ | 0.2 | 0.1 | 0.05 | 0.025 | $\nu \backslash \alpha$ | 0.2 | 0.1 | 0.05 | 0.025 |
|-------------------------|--------|--------|---------|---------|-------------------------|--------|--------|--------|--------|
| 1 | 3.0777 | 6.3138 | 12.7062 | 25.4517 | 31 | 1.3095 | 1.6955 | 2.0395 | 2.3556 |
| 2 | 1.8856 | 2.9200 | 4.3027 | 6.2053 | 32 | 1.3086 | 1.6939 | 2.0369 | 2.3518 |
| 3 | 1.6377 | 2.3534 | 3.1824 | 4.1765 | 33 | 1.3077 | 1.6924 | 2.0345 | 2.3483 |
| 4 | 1.5332 | 2.1318 | 2.7764 | 3.4954 | 34 | 1.3070 | 1.6909 | 2.0322 | 2.3451 |
| 5 | 1.4759 | 2.0150 | 2.5706 | 3.1634 | 35 | 1.3062 | 1.6896 | 2.0301 | 2.3420 |
| 6 | 1.4398 | 1.9432 | 2.4469 | 2.9687 | 36 | 1.3055 | 1.6883 | 2.0281 | 2.3391 |
| 7 | 1.4149 | 1.8946 | 2.3646 | 2.8412 | 37 | 1.3049 | 1.6871 | 2.0262 | 2.3363 |
| 8 | 1.3968 | 1.8595 | 2.3060 | 2.7515 | 38 | 1.3042 | 1.6860 | 2.0244 | 2.3337 |
| 9 | 1.3830 | 1.8331 | 2.2622 | 2.6850 | 39 | 1.3036 | 1.6849 | 2.0227 | 2.3313 |
| 10 | 1.3722 | 1.8125 | 2.2281 | 2.6338 | 40 | 1.3031 | 1.6839 | 2.0211 | 2.3289 |
| 11 | 1.3634 | 1.7959 | 2.2010 | 2.5931 | 41 | 1.3025 | 1.6829 | 2.0195 | 2.3267 |
| 12 | 1.3562 | 1.7823 | 2.1788 | 2.5600 | 42 | 1.3020 | 1.6820 | 2.0181 | 2.3246 |
| 13 | 1.3502 | 1.7709 | 2.1604 | 2.5326 | 43 | 1.3016 | 1.6811 | 2.0167 | 2.3226 |
| 14 | 1.3450 | 1.7613 | 2.1448 | 2.5096 | 44 | 1.3011 | 1.6802 | 2.0154 | 2.3207 |
| 15 | 1.3406 | 1.7531 | 2.1314 | 2.4899 | 45 | 1.3006 | 1.6794 | 2.0141 | 2.3189 |
| 16 | 1.3368 | 1.7459 | 2.1199 | 2.4729 | 46 | 1.3002 | 1.6787 | 2.0129 | 2.3172 |
| 17 | 1.3334 | 1.7396 | 2.1098 | 2.4581 | 47 | 1.2998 | 1.6779 | 2.0117 | 2.3155 |
| 18 | 1.3304 | 1.7341 | 2.1009 | 2.4450 | 48 | 1.2994 | 1.6772 | 2.0106 | 2.3139 |
| 19 | 1.3277 | 1.7291 | 2.0930 | 2.4334 | 49 | 1.2991 | 1.6766 | 2.0096 | 2.3124 |
| 20 | 1.3253 | 1.7247 | 2.0860 | 2.4231 | 50 | 1.2987 | 1.6759 | 2.0086 | 2.3109 |
| 21 | 1.3232 | 1.7207 | 2.0796 | 2.4138 | 60 | 1.2958 | 1.6706 | 2.0003 | 2.2990 |
| 22 | 1.3212 | 1.7171 | 2.0739 | 2.4055 | 70 | 1.2938 | 1.6669 | 1.9944 | 2.2906 |
| 23 | 1.3195 | 1.7139 | 2.0687 | 2.3979 | 80 | 1.2922 | 1.6641 | 1.9901 | 2.2844 |
| 24 | 1.3178 | 1.7109 | 2.0639 | 2.3909 | 90 | 1.2910 | 1.6620 | 1.9867 | 2.2795 |
| 25 | 1.3163 | 1.7081 | 2.0595 | 2.3846 | 100 | 1.2901 | 1.6602 | 1.9840 | 2.2757 |
| 26 | 1.3150 | 1.7056 | 2.0555 | 2.3788 | 110 | 1.2893 | 1.6588 | 1.9818 | 2.2725 |
| 27 | 1.3137 | 1.7033 | 2.0518 | 2.3734 | 120 | 1.2886 | 1.6577 | 1.9799 | 2.2699 |
| 28 | 1.3125 | 1.7011 | 2.0484 | 2.3685 | 130 | 1.2881 | 1.6567 | 1.9784 | 2.2677 |
| 29 | 1.3114 | 1.6991 | 2.0452 | 2.3638 | 140 | 1.2876 | 1.6558 | 1.9771 | 2.2658 |
| 30 | 1.3104 | 1.6973 | 2.0423 | 2.3596 | 150 | 1.2872 | 1.6551 | 1.9759 | 2.2641 |

Примітка. Критичні значення $t(\alpha; \nu)$ можна розрахувати у Microsoft Excel за допомогою функції `СТЮДЕНТ.ОБР(1 - 0.5 α ; ν)` категорії «Статистичні». Наприклад, для розрахунку $t(0.05; 30)$ слід задати формулу `'=СТЮДЕНТ.ОБР(1-0.025; 30)'`.

В.5 Табульовані значення статистичних критеріїв, що використовуються при обробці сукупностей з групами похибками

Табл. В.5. Критичні значення критерію Граббса

| | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.01$ | | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.01$ |
|-----------|----------------|-----------------|-----------------|-----------|----------------|-----------------|-----------------|
| 3 | 1.412 | 1.414 | 1.414 | 27 | 2.749 | 2.913 | 3.239 |
| 4 | 1.689 | 1.710 | 1.728 | 28 | 2.764 | 2.929 | 3.258 |
| 5 | 1.869 | 1.917 | 1.972 | 29 | 2.778 | 2.944 | 3.275 |
| 6 | 1.996 | 2.067 | 2.161 | 30 | 2.792 | 2.958 | 3.291 |
| 7 | 2.093 | 2.182 | 2.310 | 31 | 2.805 | 2.972 | 3.307 |
| 8 | 2.172 | 2.273 | 2.431 | 32 | 2.818 | 2.985 | 3.322 |
| 9 | 2.238 | 2.349 | 2.532 | 33 | 2.830 | 2.998 | 3.337 |
| 10 | 2.294 | 2.414 | 2.616 | 34 | 2.842 | 3.010 | 3.351 |
| 11 | 2.343 | 2.470 | 2.689 | 35 | 2.853 | 3.022 | 3.364 |
| 12 | 2.387 | 2.519 | 2.753 | 36 | 2.864 | 3.033 | 3.377 |
| 13 | 2.426 | 2.563 | 2.809 | 37 | 2.874 | 3.044 | 3.389 |
| 14 | 2.461 | 2.602 | 2.859 | 38 | 2.885 | 3.055 | 3.401 |
| 15 | 2.494 | 2.638 | 2.905 | 39 | 2.894 | 3.065 | 3.413 |
| 16 | 2.523 | 2.670 | 2.946 | 40 | 2.904 | 3.075 | 3.424 |
| 17 | 2.551 | 2.701 | 2.983 | 41 | 2.913 | 3.084 | 3.435 |
| 18 | 2.577 | 2.728 | 3.017 | 42 | 2.922 | 3.094 | 3.445 |
| 19 | 2.601 | 2.754 | 3.049 | 43 | 2.931 | 3.103 | 3.455 |
| 20 | 2.623 | 2.779 | 3.079 | 44 | 2.940 | 3.112 | 3.465 |
| 21 | 2.644 | 2.801 | 3.106 | 45 | 2.948 | 3.120 | 3.474 |
| 22 | 2.664 | 2.823 | 3.132 | 46 | 2.956 | 3.129 | 3.483 |
| 23 | 2.683 | 2.843 | 3.156 | 47 | 2.964 | 3.137 | 3.492 |
| 24 | 2.701 | 2.862 | 3.179 | 48 | 2.972 | 3.145 | 3.501 |
| 25 | 2.718 | 2.880 | 3.200 | 49 | 2.980 | 3.152 | 3.510 |
| 26 | 2.734 | 2.897 | 3.220 | 50 | 2.987 | 3.160 | 3.518 |

Табл. В.6. Критичні значення для L , L^1 -тестів Тіт'єна — Мура ($\alpha = 0.05$)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|
| 3 | 0.003 | | | | | | | | | |
| 4 | 0.051 | 0.001 | | | | | | | | |
| 5 | 0.125 | 0.018 | | | | | | | | |
| 6 | 0.203 | 0.055 | 0.010 | | | | | | | |
| 7 | 0.273 | 0.106 | 0.032 | | | | | | | |
| 8 | 0.326 | 0.146 | 0.064 | 0.022 | | | | | | |
| 9 | 0.372 | 0.194 | 0.099 | 0.045 | | | | | | |
| 10 | 0.418 | 0.233 | 0.129 | 0.070 | 0.034 | | | | | |
| 11 | 0.454 | 0.270 | 0.162 | 0.098 | 0.054 | | | | | |
| 12 | 0.489 | 0.305 | 0.196 | 0.125 | 0.076 | 0.042 | | | | |
| 13 | 0.517 | 0.337 | 0.224 | 0.150 | 0.098 | 0.060 | | | | |
| 14 | 0.540 | 0.363 | 0.250 | 0.174 | 0.122 | 0.079 | 0.055 | | | |
| 15 | 0.556 | 0.387 | 0.276 | 0.197 | 0.140 | 0.097 | 0.066 | | | |
| 16 | 0.575 | 0.410 | 0.300 | 0.219 | 0.159 | 0.115 | 0.082 | 0.055 | | |
| 17 | 0.594 | 0.427 | 0.322 | 0.240 | 0.181 | 0.136 | 0.100 | 0.072 | | |
| 18 | 0.608 | 0.447 | 0.337 | 0.259 | 0.200 | 0.154 | 0.116 | 0.086 | 0.062 | |
| 19 | 0.624 | 0.462 | 0.354 | 0.277 | 0.209 | 0.168 | 0.130 | 0.099 | 0.074 | |
| 20 | 0.639 | 0.484 | 0.377 | 0.299 | 0.238 | 0.188 | 0.150 | 0.115 | 0.088 | 0.066 |
| 25 | 0.696 | 0.550 | 0.450 | 0.374 | 0.312 | 0.262 | 0.222 | 0.184 | 0.154 | 0.126 |
| 30 | 0.730 | 0.599 | 0.506 | 0.434 | 0.376 | 0.327 | 0.283 | 0.245 | 0.212 | 0.183 |
| 35 | 0.762 | 0.642 | 0.554 | 0.482 | 0.424 | 0.376 | 0.334 | 0.297 | 0.264 | 0.235 |
| 40 | 0.784 | 0.672 | 0.588 | 0.523 | 0.468 | 0.421 | 0.378 | 0.342 | 0.310 | 0.280 |
| 45 | 0.802 | 0.696 | 0.618 | 0.556 | 0.502 | 0.456 | 0.417 | 0.382 | 0.350 | 0.320 |
| 50 | 0.820 | 0.722 | 0.646 | 0.588 | 0.535 | 0.490 | 0.450 | 0.414 | 0.384 | 0.356 |

Табл. В.7. Значення C_α -оцінки для E -тесту Тіт'єна — Мура ($\alpha = 0.05$)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|
| 3 | 0.001 | | | | | | | | | |
| 4 | 0.025 | 0.001 | | | | | | | | |
| 5 | 0.081 | 0.010 | | | | | | | | |
| 6 | 0.146 | 0.034 | 0.004 | | | | | | | |
| 7 | 0.208 | 0.065 | 0.016 | | | | | | | |
| 8 | 0.265 | 0.099 | 0.034 | 0.010 | | | | | | |
| 9 | 0.314 | 0.137 | 0.057 | 0.021 | | | | | | |
| 10 | 0.356 | 0.172 | 0.083 | 0.037 | 0.014 | | | | | |
| 11 | 0.386 | 0.204 | 0.107 | 0.055 | 0.026 | | | | | |
| 12 | 0.424 | 0.234 | 0.133 | 0.073 | 0.039 | 0.018 | | | | |
| 13 | 0.455 | 0.262 | 0.156 | 0.092 | 0.053 | 0.028 | | | | |
| 14 | 0.484 | 0.293 | 0.179 | 0.112 | 0.068 | 0.039 | 0.021 | | | |
| 15 | 0.509 | 0.317 | 0.206 | 0.134 | 0.084 | 0.052 | 0.030 | | | |
| 16 | 0.526 | 0.340 | 0.227 | 0.153 | 0.102 | 0.067 | 0.041 | 0.024 | | |
| 17 | 0.544 | 0.362 | 0.248 | 0.170 | 0.116 | 0.078 | 0.050 | 0.032 | | |
| 18 | 0.562 | 0.382 | 0.267 | 0.187 | 0.132 | 0.091 | 0.062 | 0.041 | 0.026 | |
| 19 | 0.581 | 0.398 | 0.287 | 0.203 | 0.146 | 0.105 | 0.074 | 0.050 | 0.033 | |
| 20 | 0.597 | 0.416 | 0.302 | 0.221 | 0.163 | 0.119 | 0.085 | 0.059 | 0.041 | 0.028 |
| 25 | 0.652 | 0.493 | 0.381 | 0.298 | 0.236 | 0.186 | 0.146 | 0.114 | 0.089 | 0.068 |
| 30 | 0.698 | 0.549 | 0.443 | 0.364 | 0.298 | 0.246 | 0.203 | 0.116 | 0.137 | 0.112 |
| 35 | 0.732 | 0.596 | 0.495 | 0.417 | 0.351 | 0.298 | 0.254 | 0.214 | 0.181 | 0.164 |
| 40 | 0.758 | 0.629 | 0.534 | 0.458 | 0.395 | 0.343 | 0.297 | 0.259 | 0.223 | 0.195 |
| 45 | 0.778 | 0.658 | 0.567 | 0.492 | 0.433 | 0.381 | 0.337 | 0.299 | 0.263 | 0.233 |
| 50 | 0.797 | 0.684 | 0.599 | 0.529 | 0.468 | 0.417 | 0.373 | 0.334 | 0.299 | 0.268 |

Табл. В.8. Значення C_α для E -тесту

| $k \backslash N$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 1 | 0.001 | 0.125 | 0.081 | 0.146 | 0.208 | 0.265 | 0.314 | 0.356 | 0.386 |
| 2 | | 0.001 | 0.010 | 0.034 | 0.063 | 0.099 | 0.137 | 0.172 | 0.204 |
| 3 | | | | 0.004 | 0.016 | 0.034 | 0.057 | 0.083 | 0.107 |
| $k \backslash N$ | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 1 | 0.424 | 0.455 | 0.484 | 0.509 | 0.526 | 0.544 | 0.562 | 0.581 | 0.597 |
| 2 | 0.234 | 0.262 | 0.293 | 0.317 | 0.340 | 0.362 | 0.382 | 0.397 | 0.416 |
| 3 | 0.133 | 0.165 | 0.179 | 0.206 | 0.227 | 0.248 | 0.267 | 0.287 | 0.302 |

Табл. В.9. Значення α для розрахунку стійких оцінок $T(\alpha)$ -Пуанкаре і $W(\alpha)$ -Вінзора

| ξ | α | ξ | α |
|-------|----------|-------|----------|
| 0 | 0.000 | 0.2 | 0.194 |
| 0.001 | 0.004 | 0.25 | 0.222 |
| 0.002 | 0.008 | 0.3 | 0.247 |
| 0.005 | 0.015 | 0.4 | 0.291 |
| 0.01 | 0.026 | 0.5 | 0.332 |
| 0.02 | 0.043 | 0.65 | 0.386 |
| 0.05 | 0.081 | 0.8 | 0.436 |
| 0.1 | 0.127 | 1 | 0.500 |
| 0.15 | 0.164 | | |

Табл. В.10. Значення $k = f(\xi)$ для розрахунку стійкої оцінки Хубера

| ξ | k | ξ | k |
|-------|-------|-------|-------|
| 0 | 0.000 | 0.2 | 0.862 |
| 0.001 | 2.630 | 0.25 | 0.766 |
| 0.002 | 2.435 | 0.3 | 0.685 |
| 0.005 | 2.160 | 0.4 | 0.550 |
| 0.01 | 1.945 | 0.5 | 0.436 |
| 0.02 | 1.717 | 0.65 | 0.291 |
| 0.05 | 1.399 | 0.8 | 0.162 |
| 0.1 | 1.140 | 1 | 0.000 |
| 0.15 | 0.980 | | |

В.6 Критерій Аббе

Табл. В.11. Критерій Аббе

| $\begin{array}{c c} P \\ \hline n \end{array}$ | 0.001 | 0.01 | 0.05 | $\begin{array}{c c} P \\ \hline n \end{array}$ | 0.001 | 0.01 | 0.05 |
|--|--------|--------|--------|--|--------|--------|--------|
| 4 | 0.2949 | 0.3128 | 0.3902 | 32 | 0.4963 | 0.6089 | 0.7177 |
| 5 | 0.2080 | 0.2690 | 0.4102 | 33 | 0.5027 | 0.6141 | 0.7216 |
| 6 | 0.1817 | 0.2808 | 0.4451 | 34 | 0.5090 | 0.6193 | 0.7256 |
| 7 | 0.1848 | 0.3070 | 0.4680 | 35 | 0.5150 | 0.6242 | 0.7292 |
| 8 | 0.2018 | 0.3314 | 0.4912 | 36 | 0.5208 | 0.6290 | 0.7328 |
| 9 | 0.2210 | 0.3544 | 0.5121 | 37 | 0.5265 | 0.6337 | 0.7363 |
| 10 | 0.2408 | 0.3759 | 0.5311 | 38 | 0.5319 | 0.6381 | 0.7396 |
| 11 | 0.2598 | 0.3957 | 0.5482 | 39 | 0.5373 | 0.6425 | 0.7429 |
| 12 | 0.2778 | 0.4140 | 0.5638 | 40 | 0.5425 | 0.6467 | 0.7461 |
| 13 | 0.2949 | 0.4309 | 0.5778 | 41 | 0.5475 | 0.6508 | 0.7491 |
| 14 | 0.3112 | 0.4166 | 0.5908 | 42 | 0.5524 | 0.6548 | 0.7521 |
| 15 | 0.3256 | 0.4611 | 0.6027 | 43 | 0.5571 | 0.6587 | 0.7550 |
| 16 | 0.3413 | 0.4746 | 0.6137 | 44 | 0.5616 | 0.6622 | 0.7576 |
| 17 | 0.3552 | 0.4872 | 0.6237 | 45 | 0.5660 | 0.6659 | 0.7603 |
| 18 | 0.3684 | 0.4989 | 0.6330 | 46 | 0.5701 | 0.6693 | 0.7628 |
| 19 | 0.3809 | 0.5100 | 0.6417 | 47 | 0.5743 | 0.6727 | 0.7653 |
| 20 | 0.3926 | 0.5203 | 0.6498 | 48 | 0.5781 | 0.6757 | 0.7676 |
| 21 | 0.4037 | 0.5301 | 0.6574 | 49 | 0.5817 | 0.6787 | 0.7698 |
| 22 | 0.4142 | 0.5393 | 0.6645 | 50 | 0.5853 | 0.6814 | 0.7718 |
| 23 | 0.4241 | 0.5479 | 0.6713 | 51 | 0.5887 | 0.6842 | 0.7739 |
| 24 | 0.4334 | 0.5562 | 0.6776 | 52 | 0.5922 | 0.6869 | 0.7759 |
| 25 | 0.4423 | 0.5639 | 0.6836 | 53 | 0.5955 | 0.6896 | 0.7779 |
| 26 | 0.4509 | 0.5713 | 0.6893 | 54 | 0.5989 | 0.6924 | 0.7799 |
| 27 | 0.4591 | 0.5784 | 0.6946 | 55 | 0.6020 | 0.6949 | 0.7817 |
| 28 | 0.4670 | 0.5850 | 0.6996 | 56 | 0.6051 | 0.6974 | 0.7836 |
| 29 | 0.4748 | 0.5915 | 0.7046 | 57 | 0.6083 | 0.6999 | 0.7853 |
| 30 | 0.4822 | 0.5975 | 0.7091 | 58 | 0.6114 | 0.7034 | 0.7872 |
| 31 | 0.4895 | 0.6034 | 0.7136 | 59 | 0.6145 | 0.7049 | 0.7891 |
| | | | | 60 | 0.6174 | 0.7071 | 0.7906 |

В.7 Значення середньої μ і стандартних похибок σ_1 і σ_2 для тесту Фостера — Стюарта

Табл. В.12. Значення середньої μ і стандартних похибок σ_1 і σ_2

| n | μ | σ_1 | σ_2 |
|------------|-------|------------|------------|
| 10 | 3.858 | 1.288 | 1.964 |
| 15 | 4.636 | 1.521 | 2.153 |
| 20 | 5.195 | 1.677 | 2.279 |
| 25 | 5.632 | 1.791 | 2.373 |
| 30 | 5.990 | 1.882 | 2.447 |
| 35 | 6.294 | 1.956 | 2.509 |
| 40 | 6.557 | 2.019 | 2.561 |
| 45 | 6.790 | 2.072 | 2.606 |
| 50 | 6.998 | 2.121 | 2.645 |
| 55 | 7.187 | 2.163 | 2.681 |
| 60 | 7.360 | 2.201 | 2.713 |
| 65 | 7.519 | 2.236 | 2.742 |
| 70 | 7.666 | 2.268 | 2.769 |
| 75 | 7.803 | 2.297 | 2.793 |
| 80 | 7.931 | 2.324 | 2.816 |
| 85 | 8.051 | 2.349 | 2.837 |
| 90 | 8.165 | 2.373 | 2.857 |
| 95 | 8.273 | 2.396 | 2.876 |
| 100 | 8.375 | 2.416 | 2.894 |

В.8 Статистика Дарбіна — Уотсона

Табл. В.13. Значення статистик Дарбіна — Уотсона d_0 і d_u при рівні значущості $\alpha = 0.05$ (n — кількість спостережень, p — кількість пояснюючих змінних без урахування константи)

| n | $p = 1$ | | $p = 2$ | | $p = 3$ | | $p = 4$ | | $p = 5$ | | $p = 6$ | | $p = 7$ | | $p = 8$ | | $p = 9$ | | $p = 10$ | |
|-----|---------|-------|---------|-------|---------|-------------|---------|-------------|---------|-------------|---------|-------------|---------|-------------|---------|-------------|---------|-------------|----------|-------------|
| | d_0 | d_u | d_0 | d_u | d_0 | d_u | d_0 | d_u | d_0 | d_u | d_0 | d_u | d_0 | d_u | d_0 | d_u | d_0 | d_u | d_0 | d_u |
| 6 | 0.61 | 1.40 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 7 | 0.70 | 1.36 | 0.47 | 1.90 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 8 | 0.76 | 1.33 | 0.56 | 1.78 | 0.37 | <u>2.29</u> | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 9 | 0.82 | 1.32 | 0.63 | 1.70 | 0.46 | <u>2.13</u> | 0.30 | <u>2.59</u> | — | — | — | — | — | — | — | — | — | — | — | — |
| 10 | 0.88 | 1.32 | 0.70 | 1.64 | 0.53 | <u>2.02</u> | 0.38 | <u>2.41</u> | 0.24 | <u>2.82</u> | — | — | — | — | — | — | — | — | — | — |
| 11 | 0.93 | 1.32 | 0.66 | 1.60 | 0.60 | 1.93 | 0.44 | <u>2.28</u> | 0.32 | <u>2.65</u> | 0.20 | <u>3.01</u> | — | — | — | — | — | — | — | — |
| 12 | 0.97 | 1.33 | 0.81 | 1.58 | 0.66 | 1.86 | 0.51 | <u>2.18</u> | 0.38 | <u>2.51</u> | 0.27 | <u>2.83</u> | 0.17 | <u>3.15</u> | — | — | — | — | — | — |
| 13 | 1.01 | 1.34 | 0.86 | 1.56 | 0.72 | 1.82 | 0.57 | <u>2.09</u> | 0.45 | <u>2.39</u> | 0.33 | <u>2.69</u> | 0.23 | <u>2.99</u> | 0.15 | <u>3.27</u> | — | — | — | — |
| 14 | 1.05 | 1.35 | 0.91 | 1.55 | 0.77 | 1.78 | 0.63 | <u>2.03</u> | 0.51 | <u>2.30</u> | 0.39 | <u>2.57</u> | 0.29 | <u>2.85</u> | 0.20 | <u>3.11</u> | 0.13 | <u>3.36</u> | — | — |
| 15 | 1.08 | 1.36 | 0.95 | 1.54 | 0.81 | 1.75 | 0.69 | 1.98 | 0.56 | <u>2.22</u> | 0.45 | <u>2.47</u> | 0.34 | <u>2.73</u> | 0.25 | <u>2.98</u> | 0.18 | <u>3.22</u> | 0.11 | <u>3.44</u> |
| 16 | 1.11 | 1.37 | 0.98 | 1.54 | 0.86 | 1.73 | 0.73 | 1.94 | 0.62 | <u>2.16</u> | 0.50 | <u>2.39</u> | 0.40 | <u>2.62</u> | 0.30 | <u>2.86</u> | 0.22 | <u>3.09</u> | 0.16 | <u>3.30</u> |
| 17 | 1.13 | 1.38 | 1.02 | 1.54 | 0.90 | 1.71 | 0.78 | 1.90 | 0.66 | <u>2.10</u> | 0.55 | <u>2.32</u> | 0.45 | <u>2.54</u> | 0.36 | <u>2.76</u> | 0.27 | <u>2.98</u> | 0.20 | <u>3.18</u> |
| 18 | 1.16 | 1.39 | 1.05 | 1.54 | 0.93 | 1.70 | 0.82 | 1.87 | 0.71 | <u>2.06</u> | 0.60 | <u>2.26</u> | 0.50 | <u>2.46</u> | 0.41 | <u>2.67</u> | 0.32 | <u>2.87</u> | 0.24 | <u>3.07</u> |
| 19 | 1.18 | 1.40 | 1.07 | 1.54 | 0.97 | 1.69 | 0.86 | 1.85 | 0.75 | <u>2.02</u> | 0.65 | <u>2.21</u> | 0.55 | <u>2.40</u> | 0.46 | <u>2.59</u> | 0.37 | <u>2.78</u> | 0.29 | <u>2.97</u> |
| 20 | 1.20 | 1.41 | 1.10 | 1.54 | 1.00 | 1.68 | 0.98 | 1.83 | 0.79 | 1.99 | 0.69 | <u>2.16</u> | 0.60 | <u>2.34</u> | 0.50 | <u>2.52</u> | 0.42 | <u>2.70</u> | 0.34 | <u>2.89</u> |
| 21 | 1.22 | 1.42 | 1.13 | 1.54 | 1.03 | 1.67 | 0.93 | 1.81 | 0.83 | 1.96 | 0.73 | <u>2.12</u> | 0.64 | <u>2.29</u> | 0.55 | <u>2.46</u> | 0.46 | <u>2.63</u> | 0.38 | <u>2.81</u> |
| 22 | 1.24 | 1.43 | 1.15 | 1.54 | 1.05 | 1.66 | 0.96 | 1.80 | 0.86 | 1.94 | 0.77 | <u>2.09</u> | 0.68 | <u>2.25</u> | 0.59 | <u>2.41</u> | 0.50 | <u>2.57</u> | 0.42 | <u>2.73</u> |
| 23 | 1.26 | 1.44 | 1.17 | 1.54 | 1.08 | 1.66 | 0.99 | 1.79 | 0.90 | 1.92 | 0.80 | <u>2.06</u> | 0.72 | <u>2.21</u> | 0.63 | <u>2.36</u> | 0.55 | <u>2.51</u> | 0.47 | <u>2.67</u> |
| 24 | 1.27 | 1.45 | 1.19 | 1.55 | 1.10 | 1.66 | 1.01 | 1.78 | 0.93 | 1.90 | 0.84 | <u>2.04</u> | 0.75 | <u>2.17</u> | 0.67 | <u>2.32</u> | 0.58 | <u>2.46</u> | 0.51 | <u>2.61</u> |
| 25 | 1.29 | 1.45 | 1.21 | 1.55 | 1.12 | 1.65 | 1.04 | 1.77 | 0.95 | 1.89 | 0.87 | <u>2.01</u> | 0.78 | <u>2.14</u> | 0.70 | <u>2.28</u> | 0.62 | <u>2.42</u> | 0.54 | <u>2.56</u> |
| 26 | 1.30 | 1.46 | 1.22 | 1.55 | 1.14 | 1.65 | 1.06 | 1.76 | 0.98 | 1.87 | 0.90 | 1.99 | 0.82 | <u>2.12</u> | 0.74 | <u>2.25</u> | 0.66 | <u>2.38</u> | 0.58 | <u>2.51</u> |
| 27 | 1.32 | 1.47 | 1.24 | 1.56 | 1.16 | 1.65 | 1.08 | 1.75 | 1.00 | 1.86 | 0.93 | 1.97 | 0.85 | <u>2.09</u> | 0.77 | <u>2.22</u> | 0.69 | <u>2.34</u> | 0.62 | <u>2.47</u> |
| 28 | 1.33 | 1.48 | 1.26 | 1.56 | 1.18 | 1.65 | 1.10 | 1.75 | 1.03 | 1.85 | 0.95 | 1.96 | 0.87 | <u>2.07</u> | 0.80 | <u>2.19</u> | 0.72 | <u>2.31</u> | 0.65 | <u>2.43</u> |
| 29 | 1.34 | 1.48 | 1.27 | 1.56 | 1.20 | 1.65 | 1.12 | 1.74 | 1.05 | 1.84 | 0.98 | 1.94 | 0.90 | <u>2.05</u> | 0.83 | <u>2.16</u> | 0.75 | <u>2.28</u> | 0.68 | <u>2.40</u> |
| 30 | 1.35 | 1.49 | 1.28 | 1.57 | 1.21 | 1.65 | 1.14 | 1.74 | 1.07 | 1.83 | 1.00 | 1.93 | 0.93 | <u>2.03</u> | 0.85 | <u>2.14</u> | 0.78 | <u>2.25</u> | 0.71 | <u>2.36</u> |
| 31 | 1.36 | 1.50 | 1.30 | 1.57 | 1.23 | 1.65 | 1.16 | 1.74 | 1.09 | 1.83 | 1.02 | 1.92 | 0.95 | <u>2.02</u> | 0.88 | <u>2.12</u> | 0.81 | <u>2.23</u> | 0.74 | <u>2.33</u> |
| 32 | 1.37 | 1.50 | 1.31 | 1.57 | 1.24 | 1.65 | 1.18 | 1.73 | 1.11 | 1.82 | 1.04 | 1.91 | 0.97 | <u>2.00</u> | 0.90 | <u>2.10</u> | 0.84 | <u>2.20</u> | 0.77 | <u>2.31</u> |
| 33 | 1.38 | 1.51 | 1.32 | 1.58 | 1.26 | 1.65 | 1.19 | 1.73 | 1.13 | 1.81 | 1.06 | 1.90 | 0.99 | 1.99 | 0.93 | <u>2.09</u> | 0.86 | <u>2.18</u> | 0.80 | <u>2.28</u> |
| 34 | 1.39 | 1.51 | 1.33 | 1.58 | 1.27 | 1.65 | 1.03 | 1.73 | 1.14 | 1.81 | 1.08 | 1.89 | 1.02 | 1.98 | 0.95 | <u>2.07</u> | 0.89 | <u>2.16</u> | 0.82 | <u>2.26</u> |
| 35 | 1.40 | 1.52 | 1.34 | 1.58 | 1.28 | 1.65 | 1.22 | 1.73 | 1.16 | 1.80 | 1.10 | 1.88 | 1.03 | 1.97 | 0.97 | <u>2.05</u> | 0.91 | <u>2.14</u> | 0.85 | <u>2.24</u> |
| 36 | 1.41 | 1.53 | 1.35 | 1.59 | 1.30 | 1.65 | 1.24 | 1.72 | 1.18 | 1.80 | 1.11 | 1.88 | 1.05 | 1.96 | 0.99 | <u>2.04</u> | 0.93 | <u>2.13</u> | 0.87 | <u>2.22</u> |
| 37 | 1.42 | 1.53 | 1.36 | 1.59 | 1.31 | 1.66 | 1.25 | 1.72 | 1.19 | 1.80 | 1.13 | 1.87 | 1.07 | 1.95 | 1.01 | <u>2.03</u> | 0.95 | <u>2.11</u> | 0.89 | <u>2.20</u> |
| 38 | 1.43 | 1.54 | 1.37 | 1.59 | 1.32 | 1.66 | 1.26 | 1.72 | 1.20 | 1.79 | 1.15 | 1.86 | 1.09 | 1.94 | 1.03 | <u>2.02</u> | 0.97 | <u>2.10</u> | 0.91 | <u>2.18</u> |
| 39 | 1.44 | 1.54 | 1.38 | 1.60 | 1.33 | 1.66 | 1.27 | 1.72 | 1.22 | 1.79 | 1.16 | 1.86 | 1.10 | 1.93 | 1.05 | <u>2.01</u> | 0.99 | <u>2.09</u> | 0.93 | <u>2.16</u> |
| 40 | 1.44 | 1.54 | 1.39 | 1.60 | 1.34 | 1.66 | 1.29 | 1.72 | 1.23 | 1.79 | 1.18 | 1.85 | 1.12 | 1.92 | 1.06 | 2.00 | 1.01 | <u>2.07</u> | 0.95 | <u>2.15</u> |
| 45 | 1.48 | 1.57 | 1.43 | 1.62 | 1.38 | 1.67 | 1.34 | 1.72 | 1.29 | 1.78 | 1.24 | 1.84 | 1.19 | 1.90 | 1.14 | 1.96 | 1.09 | <u>2.02</u> | 1.04 | <u>2.09</u> |
| 50 | 1.50 | 1.59 | 1.46 | 1.63 | 1.42 | 1.67 | 1.38 | 1.72 | 1.34 | 1.77 | 1.29 | 1.82 | 1.25 | 1.88 | 1.20 | 1.93 | 1.16 | 1.99 | 1.11 | <u>2.04</u> |
| 55 | 1.53 | 1.60 | 1.49 | 1.64 | 1.45 | 1.68 | 1.41 | 1.72 | 1.37 | 1.77 | 1.33 | 1.81 | 1.29 | 1.86 | 1.25 | 1.91 | 1.21 | 1.96 | 1.17 | <u>2.01</u> |
| 60 | 1.55 | 1.62 | 1.51 | 1.65 | 1.48 | 1.69 | 1.44 | 1.73 | 1.41 | 1.77 | 1.37 | 1.81 | 1.34 | 1.85 | 1.30 | 1.89 | 1.26 | 1.94 | 1.22 | 1.98 |
| 65 | 1.57 | 1.63 | 1.54 | 1.66 | 1.50 | 1.70 | 1.47 | 1.73 | 1.44 | 1.77 | 1.40 | 1.81 | 1.37 | 1.84 | 1.34 | 1.88 | 1.30 | 1.92 | 1.27 | 1.96 |
| 70 | 1.58 | 1.64 | 1.55 | 1.67 | 1.53 | 1.70 | 1.49 | 1.74 | 1.46 | 1.77 | 1.43 | 1.80 | 1.40 | 1.84 | 1.37 | 1.87 | 1.34 | 1.91 | 1.31 | 1.95 |
| 75 | 1.60 | 1.65 | 1.57 | 1.68 | 1.54 | 1.71 | 1.52 | 1.74 | 1.49 | 1.77 | 1.46 | 1.80 | 1.43 | 1.83 | 1.40 | 1.87 | 1.37 | 1.90 | 1.34 | 1.94 |
| 80 | 1.61 | 1.66 | 1.59 | 1.69 | 1.56 | 1.72 | 1.53 | 1.74 | 1.51 | 1.77 | 1.48 | 1.80 | 1.45 | 1.83 | 1.43 | 1.86 | 1.40 | 1.89 | 1.37 | 1.93 |
| 85 | 1.62 | 1.67 | 1.60 | 1.70 | 1.58 | 1.72 | 1.55 | 1.75 | 1.53 | 1.77 | 1.50 | 1.80 | 1.47 | 1.83 | 1.45 | 1.86 | 1.42 | 1.89 | 1.40 | 1.92 |
| 90 | 1.64 | 1.68 | 1.61 | 1.70 | 1.59 | 1.73 | 1.57 | 1.75 | 1.54 | 1.78 | 1.52 | 1.80 | 1.49 | 1.83 | 1.47 | 1.85 | 1.45 | 1.88 | 1.42 | 1.91 |
| 95 | 1.65 | 1.69 | 1.62 | 1.71 | 1.60 | 1.73 | 1.58 | 1.76 | 1.56 | 1.78 | 1.54 | 1.80 | 1.51 | 1.83 | 1.49 | 1.85 | 1.47 | 1.88 | 1.44 | 1.90 |
| 100 | 1.65 | 1.69 | 1.63 | 1.72 | 1.61 | 1.74 | 1.59 | 1.76 | 1.57 | 1.78 | 1.55 | 1.80 | 1.53 | 1.83 | 1.51 | 1.85 | 1.48 | 1.87 | 1.46 | 1.90 |
| 150 | 1.72 | 1.75 | 1.71 | 1.76 | 1.69 | 1.77 | 1.68 | 1.79 | 1.67 | 1.80 | 1.65 | 1.82 | 1.64 | 1.83 | 1.62 | 1.85 | 1.61 | 1.86 | 1.59 | 1.88 |
| 200 | 1.76 | 1.78 | 1.75 | 1.79 | 1.74 | 1.80 | 1.73 | 1.81 | 1.72 | 1.82 | 1.71 | 1.83 | 1.70 | 1.84 | 1.69 | 1.85 | 1.68 | 1.86 | 1.67 | 1.87 |

Примітка. Підкреслені значення d_u означають неможливість отримання оцінки відсутності автокореляції залишків за допомогою даного критерію при відповідних n і p .

В.9 Критерій Девіда — Хартлі — Пірсона (RS-критерій)

Табл. В.14. Критичні межі $U_1(\alpha)$ і $U_2(\alpha)$ критерію нормальності розподілу Девіда — Хартлі — Пірсона (RS-критерію)

| | $\alpha = 0.01$ | | $\alpha = 0.05$ | | $\alpha = 0.10$ | |
|------------|-----------------|-------|-----------------|-------|-----------------|-------|
| | U_1 | U_2 | U_1 | U_2 | U_1 | U_2 |
| 3 | 1.737 | 2.000 | 1.758 | 1.999 | 1.782 | 1.997 |
| 4 | 1.870 | 2.445 | 1.980 | 2.429 | 2.040 | 2.409 |
| 5 | 2.020 | 2.803 | 2.150 | 2.753 | 2.220 | 2.712 |
| 6 | 2.150 | 3.095 | 2.200 | 3.012 | 2.370 | 2.949 |
| 7 | 2.260 | 3.338 | 2.400 | 3.222 | 2.490 | 3.143 |
| 8 | 2.350 | 3.543 | 2.500 | 3.399 | 2.590 | 3.308 |
| 9 | 2.440 | 3.720 | 2.590 | 3.552 | 2.680 | 3.449 |
| 10 | 2.510 | 3.875 | 2.670 | 3.685 | 2.760 | 3.570 |
| 11 | 2.580 | 4.012 | 2.740 | 3.800 | 2.840 | 3.680 |
| 12 | 2.640 | 4.134 | 2.800 | 3.910 | 2.900 | 3.780 |
| 13 | 2.700 | 4.244 | 2.860 | 4.000 | 2.960 | 3.870 |
| 14 | 2.750 | 4.340 | 2.920 | 4.090 | 3.020 | 3.950 |
| 15 | 2.800 | 4.440 | 2.970 | 4.170 | 3.070 | 4.020 |
| 16 | 2.840 | 4.520 | 3.010 | 4.240 | 3.120 | 4.090 |
| 17 | 2.880 | 4.600 | 3.060 | 4.310 | 3.170 | 4.150 |
| 18 | 2.920 | 4.670 | 3.100 | 4.370 | 3.210 | 4.210 |
| 19 | 2.960 | 4.740 | 3.140 | 4.430 | 3.250 | 4.270 |
| 20 | 2.990 | 4.800 | 3.180 | 4.490 | 3.290 | 4.320 |
| 25 | 3.150 | 5.060 | 3.340 | 4.710 | 3.450 | 4.530 |
| 30 | 3.270 | 5.260 | 3.470 | 4.890 | 3.590 | 4.700 |
| 35 | 3.380 | 5.420 | 3.580 | 5.040 | 3.700 | 4.840 |
| 40 | 3.470 | 5.560 | 3.670 | 5.160 | 3.790 | 4.960 |
| 45 | 3.550 | 5.670 | 3.750 | 5.260 | 3.880 | 5.060 |
| 50 | 3.620 | 5.770 | 3.830 | 5.350 | 3.950 | 5.140 |
| 55 | 3.690 | 5.860 | 3.900 | 5.430 | 4.020 | 5.220 |
| 60 | 3.750 | 5.940 | 3.960 | 5.510 | 4.080 | 5.290 |
| 65 | 3.800 | 6.010 | 4.010 | 5.570 | 4.140 | 5.350 |
| 70 | 3.850 | 6.070 | 4.060 | 5.630 | 4.190 | 5.410 |
| 75 | 3.900 | 6.130 | 4.130 | 5.680 | 4.240 | 5.460 |
| 80 | 3.940 | 6.180 | 4.150 | 5.730 | 4.280 | 5.510 |
| 85 | 3.990 | 6.230 | 4.200 | 5.780 | 4.330 | 5.560 |
| 90 | 4.020 | 6.270 | 4.240 | 5.820 | 4.360 | 5.600 |
| 95 | 4.060 | 6.320 | 4.270 | 5.860 | 4.400 | 5.640 |
| 100 | 4.100 | 6.360 | 4.310 | 5.900 | 4.440 | 5.680 |

В.10 Критерій Фроціні

Табл. В.15. Критичні значення статистики Фроціні B_N для перевірки нормальності розподілу

| N | Рівень значущості α | | | | |
|-----------|----------------------------|--------|--------|--------|--------|
| | 0.20 | 0.15 | 0.10 | 0.05 | 0.01 |
| 5 | 0.2115 | 0.2245 | 0.2422 | 0.2666 | 0.3120 |
| 6 | 0.2132 | 0.2265 | 0.2434 | 0.2698 | 0.3148 |
| 7 | 0.2143 | 0.2276 | 0.2448 | 0.2702 | 0.3224 |
| 8 | 0.2175 | 0.2303 | 0.2477 | 0.2756 | 0.3286 |
| 9 | 0.2186 | 0.2314 | 0.2491 | 0.2753 | 0.3331 |
| 10 | 0.2179 | 0.2304 | 0.2485 | 0.2789 | 0.3332 |
| 11 | 0.2191 | 0.2332 | 0.2500 | 0.2774 | 0.3317 |
| 12 | 0.2212 | 0.2335 | 0.2508 | 0.2795 | 0.3356 |
| 13 | 0.2207 | 0.2336 | 0.2525 | 0.2784 | 0.3385 |
| 14 | 0.2204 | 0.2335 | 0.2518 | 0.2791 | 0.3367 |
| 15 | 0.2226 | 0.2354 | 0.2530 | 0.2820 | 0.3370 |
| 16 | 0.2226 | 0.2356 | 0.2519 | 0.2804 | 0.3376 |
| 17 | 0.2228 | 0.2362 | 0.2551 | 0.2812 | 0.3376 |
| 18 | 0.2233 | 0.2363 | 0.2536 | 0.2822 | 0.3374 |
| 19 | 0.2232 | 0.2367 | 0.2533 | 0.2830 | 0.3413 |
| 20 | 0.2257 | 0.2385 | 0.2556 | 0.2839 | 0.3363 |
| ∞ | 0.2250 | 0.2390 | 0.2560 | 0.2840 | 0.3410 |

В.11 Стандартний нормальний розподіл

Табл. В.16. Значення функції $\Phi(x) = \Phi(x; 0; 1)$ стандартного нормального розподілу

| x | $\Phi(x)$ | x | $\Phi(x)$ | x | $\Phi(x)$ |
|-------------|-----------|-------------|-----------|-------------|-----------|
| 0.00 | 0.500000 | 1.00 | 0.841345 | 2.00 | 0.977250 |
| 0.05 | 0.519939 | 1.05 | 0.853141 | 2.05 | 0.979818 |
| 0.10 | 0.539828 | 1.10 | 0.864334 | 2.10 | 0.982136 |
| 0.15 | 0.559618 | 1.15 | 0.874928 | 2.15 | 0.984222 |
| 0.20 | 0.579260 | 1.20 | 0.884930 | 2.20 | 0.986097 |
| 0.25 | 0.598706 | 1.25 | 0.894350 | 2.25 | 0.987776 |
| 0.30 | 0.617911 | 1.30 | 0.903200 | 2.30 | 0.989276 |
| 0.35 | 0.636831 | 1.35 | 0.911492 | 2.35 | 0.990613 |
| 0.40 | 0.655422 | 1.40 | 0.919243 | 2.40 | 0.991802 |
| 0.45 | 0.673645 | 1.45 | 0.926471 | 2.45 | 0.992857 |
| 0.50 | 0.691462 | 1.50 | 0.933193 | 2.50 | 0.993790 |
| 0.55 | 0.708840 | 1.55 | 0.939429 | 2.55 | 0.994614 |
| 0.60 | 0.725747 | 1.60 | 0.945201 | 2.60 | 0.995339 |
| 0.65 | 0.742154 | 1.65 | 0.950529 | 2.65 | 0.995975 |
| 0.70 | 0.758036 | 1.70 | 0.955435 | 2.70 | 0.996533 |
| 0.75 | 0.773373 | 1.75 | 0.959941 | 2.75 | 0.997020 |
| 0.80 | 0.788145 | 1.80 | 0.964070 | 2.80 | 0.997445 |
| 0.85 | 0.802337 | 1.85 | 0.967843 | 2.85 | 0.997814 |
| 0.90 | 0.815940 | 1.90 | 0.971283 | 2.90 | 0.998134 |
| 0.95 | 0.828944 | 1.95 | 0.974412 | 2.95 | 0.998411 |
| | | | | 3.00 | 0.998650 |

Примітки.

1. Якщо величина аргументу попадає між табличними значеннями x , то для визначення $\Phi(x)$ користуються лінійною інтерполяцією функції $\Phi(x)$.
2. При визначенні значення функції $\Phi(x)$ для від'ємних величин x необхідно скористатися тотожністю $\Phi(x) = 1 - \Phi(-x)$.
3. Значення $\Phi(x)$ можна розрахувати у Microsoft Excel за допомогою функції `НОРМ.РАСП(x ; 0; 1; 1)` категорії «Статистичні». Наприклад, щоб розрахувати $\Phi(2.55)$, слід задати формулу `'=НОРМ.РАСП(2.55; 0; 1; 1)'`.