## **Автокореляція**. **Статистичні критерії перевірки вибірки на автокореляцію**

1Сутність явища автокореляції залишків та її наслідки

2Перевірка на автокореляцію за допомогою критерію знаків

3Перевірка на автокореляцію за допомогою критерію Дарбіна-Уотсона

### 1Сутність явища автокореляції залишків та її наслідки

Одне з припущень|гадка| класичного регресійного аналізу – незалежність помилок еi . Якщо це припущення|гадка| порушується, то має місце автокореляція залишків, яка пов’язана з|із| інерційністю і циклічністю економічних процесів. Найчастіше вона спостерігається|походити| при аналізі часових рядів і призводить до наступних негативних наслідків: оцінки параметрів стають неефективними, внаслідок великих значень |цебто| дисперсії для параметрів математичного очікування|чекання|, стандартні помилки коефіцієнтів стають заниженими, внаслідок чого при аналізі некоректно використовувати t і F-критерій, дисперсія прогнозних значень достатньо велика.

Дослідження залишків дозволяє зробити висновок про правильність моделі і її придатність для прогнозування. Для виявлення автокореляції застосовують як графічний аналіз залишків, так і низку статистичних критеріїв.

Аналіз графічного зображення залишків ei  включає наступні етапи:

* побудова|шикування| графіка часової послідовності, якщо залишки отримані|одержані| в різні моменти часу;
* побудова|шикування| графіка залежності ei від  (розрахунковим);
* побудова|шикування| графіка залежності ei від пояснюючих змінних x1,x2.

Якщо в результаті аналізу отримано горизонтальну лінію , можна зробити висновок щодо відсутності явища автокореляції, інакше можна говорити про неадекватність моделі або помилковість розрахунків, або необхідність включення в модель залежності від часу лінійного або квадратичного члена, або непостійність|незмінність| дисперсії тощо.

Виділяють серіальну кореляцію (автокореляцію), коли залежність між помилками, що відстають на деяке число кроків S ( порядок|лад| кореляції (в т.ч. S=1), залишається однаковою.

### 2 Перевірка на автокореляцію за допомогою критерію знаків

Одним із статистичний критерій перевірки існування автокореляції є критерій знаків. Для застосування цього критерію потрібно розташувати залишки ei в часовій послідовності, виписати їх знаки, підрахувати|підсумувати| число серій з|із| однаковими знаками, що утворились, n4, кількість залишків зі знаком плюс n1, |із| кількість залишків зі знаком мінус n2.|із|Далі визначається ймовірність появи n4 группри нульовій гіпотезі – автокореляція відсутня. Якщо, то нульова гіпотеза відкидається, тобто приймається гіпотеза щодо існування автокореляції.

Для вибірок з|із| n1, n2 ≤20 існує таблиця нормального розподілу з|із| критичними значеннями n4 при q=0,05.

Для великих вибірок розподіл помилок апроксимується нормальним законом з|із| середнім  і дисперсією , а величина  підпорядковується нормованому стандартному нормальному розподілу. Критичні значення n4 знаходять|находити|за формулами , де zα визначається за|із| умови  , де знаходиться в|із| таблицях нормального розподілу.

Отже, для перевірки вибірки на автокореляцію за критерієм знаків необхідно:

* знайти параметри моделі  та ;
* побудувати регресійну модель;

– знайти залишки регресійної моделі  і записати їх за часовою послідовністю;

* виписати знаки (додатні або від’ємні) залишків;
* порахувати кількість серій  з однаковими знаками;
* порахувати кількість  додатних залишків;
* порахувати кількість  від’ємних залишків;
* за статистичними таблицями знайти верхню та нижню межі для .

Нижня межа  розташована у верхній частині таблиці на перехресті рядка  та стовпчика . Верхня межа  розташована у нижній частині таблиці. Якщо розрахункове значення  знаходиться в отриманих межах, автокореляція відсутня.

### 3 Перевірка вибірки на автокореляцію за допомогою критерію Дарбіна-Уотсона

Дарбін і Уотсон запропонували оцінювати автокореляцію за допомогою статистики:



або

;

де  – вибірковий коефіцієнт автокореляції першого порядку|лад|.

При n→∞ (n>100) можна нехтувати доданком  і статистика . Область значень . При сильній позитивній автокореляції і , при сильній негативній|заперечний| автокореляції  і , за відсутності автокореляції , .

Для статистики d є|наявний| верхній (upper) du і нижній (low) dl межі рівня значущості. Ці значення для відомих значень α (або q), n – обсягу вибірки і k– розмірності моделі, табульовані.

Для перевірки вибірки на автокореляцію за критерієм Дарбіна-Уотсона розрахуємо статистику d:

.

Для статистики  верхня  та нижня  межі рівня значущості при *k* незалежних змінних моделі табульовані.

При додатній автокореляції область розрахункових значень  відповідає гіпотезі Н0 (відсутність автокореляції з імовірністю ), а область – гіпотезі *Н1*(наявність автокореляції з імовірністю ). Проміжні значення називають областю невизначеності.

Тест застосовується для додатної та від’ємної автокореляцій, що спостерігаються у разі частої зміни знаків *ei*. В цьому випадку область  відповідає гіпотезі *Н0*, область  – випадку невизначеності та  – гіпотезі *Н1* (табл.6.1).

Таблиця 1.12 – Області наявності або відсутності автокореляції

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *d* | () |  |  |  |  |
| *Hi* | *H1* | *?* | *H0* | *?* | *H1* |

Наприклад, якщо розрахункове значення  потрапляє в область   
2<<(4–), можна зробити висновок про відсутність автокореляції.

**Приклад 1.18**

Перевірити вибірку на автокореляцію критеріями знаків та Дарбіна-Уотсона (табл.1.13).

Таблиця 1.13 – Вибірка до прикладу 1.18.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Рік | X | Y | Рік | X | Y | Рік | X | Y |
| 1990 | 1 | 40,69 | 1992 | 9 | 110,4 | 1994 | 17 | 248,7 |
| 1990 | 2 | 68,76 | 1992 | 10 | 184,2 | 1994 | 18 | 405,3 |
| 1990 | 3 | 91,76 | 1992 | 11 | 240,1 | 1994 | 19 | 514,9 |
| 1990 | 4 | 45,03 | 1992 | 12 | 114,2 | 1994 | 20 | 240,6 |
| 1991 | 5 | 67,69 | 1993 | 13 | 168,8 | 1995 | 21 | 347,9 |
| 1991 | 6 | 115,9 | 1993 | 14 | 279,6 | 1995 | 22 | 562 |
| 1991 | 7 | 152,7 | 1993 | 15 | 359,6 | 1995 | 23 | 708,2 |
| 1991 | 8 | 73,84 | 1993 | 16 | 171 | 1995 | 24 | 327,3 |

**Розв’язання**

Знайдемо параметри моделей  та  за допомогою вбудованої функції Excel ЛИНЕЙН (рис. 6.1): = –24,33; =20,74.

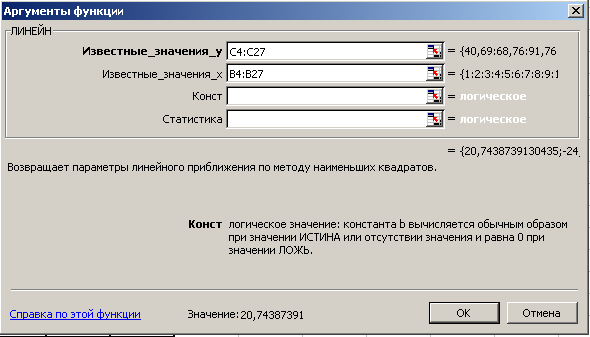


Рисунок 1.15 – ­­Структура вбудованої функції Excel ЛИНЕЙН

Знайдемо залишки регресійної моделі, яка має вигляд , та випишемо їх знаки (табл. 1.14).

Підрахуємо кількість додатних залишків , від’ємних  та кількість серій знаків . (=11, =13, =12). За статистичними таблицями знайдемо верхню та нижню межі для . Вона дорівнює 12.

Таблиця 1.14 – Результати розрахунків до прикладу 6 (за критерієм знаків)

| Рік | X | Y | Ŷ | Залишки  *еі* | Знаки залишків |
| --- | --- | --- | --- | --- | --- |
| 1990 | 1 | 40,69 | –3,59 | –44,28 | – |
| 1990 | 2 | 68,76 | 17,15 | –51,61 | – |
| 1990 | 3 | 91,76 | 37,90 | –53,86 | – |
| 1990 | 4 | 45,03 | 58,64 | 13,61 | + |
| 1991 | 5 | 67,69 | 79,39 | 11,70 | + |
| 1991 | 6 | 115,9 | 100,13 | –15,77 | – |
| 1991 | 7 | 152,7 | 120,87 | –31,83 | – |
| 1991 | 8 | 73,84 | 141,62 | 67,78 | + |
| 1992 | 9 | 110,4 | 162,36 | 51,96 | + |
| 1992 | 10 | 184,2 | 183,11 | –1,09 | – |
| 1992 | 11 | 240,1 | 203,85 | –36,25 | – |
| 1992 | 12 | 114,2 | 224,59 | 110,39 | + |
| 1993 | 13 | 168,8 | 245,34 | 76,54 | + |
| 1993 | 14 | 279,6 | 266,08 | –13,52 | – |
| 1993 | 15 | 359,6 | 286,83 | –72,77 | – |
| 1993 | 16 | 171 | 307,57 | 136,57 | + |
| 1994 | 17 | 248,7 | 328,31 | 79,61 | + |
| 1994 | 18 | 405,3 | 349,06 | –56,24 | – |
| 1994 | 19 | 514,9 | 369,80 | –145,10 | – |
| 1994 | 20 | 240,6 | 390,54 | 149,94 | + |
| 1995 | 21 | 347,9 | 411,29 | 63,39 | + |
| 1995 | 22 | 562 | 432,03 | –129,97 | – |
| 1995 | 23 | 708,2 | 452,78 | –255,42 | – |
| 1995 | 24 | 327,3 | 473,52 | 146,22 | + |

У верхній частині таблиці знаходимо нижню межу. Для цього обираємо необхідний нам рядок =11 та стовпчик =12 і отримуємо значення 7. У нижній частині таблиці знайдемо верхню межу для тих самих значень  та . Верхня межа дорівнює 17. Значення  знаходиться в отриманих межах: 7<<17; 7<12<17. Можна зробити висновок, що за критерієм знаків автокореляція відсутня.

Для перевірки вибірки на наявність автокореляції за критерієм Дарбіна-Уотсона необхідно заповнити відповідну таблицю (табл. 1.15):

Таблиця 1.15 – Результати розрахунків до прикладу 1.18 (за критерієм Дарбіна-Уотсона)

| Рік | Xі | Yі | Ŷі | еі | еі2 | еі –1 | еі–1 – еі | (еі–1 – еі)2 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1990 | 1 | 40,69 | –3,59 | –44,28 | 1960,642 |  |  |  |
| 1990 | 2 | 68,76 | 17,15 | –51,61 | 2663,103 | –44,28 | –7,33 | 53,67212 |
| 1990 | 3 | 91,76 | 37,90 | –53,86 | 2901,049 | –51,61 | –2,26 | 5,090105 |
| 1990 | 4 | 45,03 | 58,64 | 13,61 | 185,2998 | –53,86 | 67,47 | 4552,724 |
| 1991 | 5 | 67,69 | 79,39 | 11,70 | 136,8049 | 13,61 | –1,92 | 3,671539 |
| 1991 | 6 | 115,9 | 100,13 | –15,77 | 248,6854 | 11,70 | –27,47 | 754,3881 |
| 1991 | 7 | 152,7 | 120,87 | –31,83 | 1012,887 | –15,77 | –16,06 | 257,7992 |
| 1991 | 8 | 73,84 | 141,62 | 67,78 | 4593,855 | –31,83 | 99,60 | 9920,932 |
| 1992 | 9 | 110,4 | 162,36 | 51,96 | 2700,035 | 67,78 | –15,82 | 250,1498 |
| 1992 | 10 | 184,2 | 183,11 | –1,09 | 1,197423 | 51,96 | –53,06 | 2814,953 |
| 1992 | 11 | 240,1 | 203,85 | –36,25 | 1314,091 | –1,09 | –35,16 | 1235,953 |
| 1992 | 12 | 114,2 | 224,59 | 110,39 | 12186,72 | –36,25 | 146,64 | 21504,43 |
| 1993 | 13 | 168,8 | 245,34 | 76,54 | 5857,966 | 110,39 | –33,86 | 1146,237 |
| 1993 | 14 | 279,6 | 266,08 | –13,52 | 182,7572 | 76,54 | –90,06 | 8110,106 |
| 1993 | 15 | 359,6 | 286,83 | –72,77 | 5296,186 | –13,52 | –59,26 | 3511,288 |
| 1993 | 16 | 171 | 307,57 | 136,57 | 18651,09 | –72,77 | 209,34 | 43824,86 |
| 1994 | 17 | 248,7 | 328,31 | 79,61 | 6338,206 | 136,57 | –56,96 | 3244 |
| 1994 | 18 | 405,3 | 349,06 | –56,24 | 3163,306 | 79,61 | –135,86 | 18456,89 |
| 1994 | 19 | 514,9 | 369,80 | –145,10 | 21053,84 | –56,24 | –88,86 | 7895,411 |
| 1994 | 20 | 240,6 | 390,54 | 149,94 | 22483,34 | –145,10 | 295,04 | 87050,89 |
| 1995 | 21 | 347,9 | 411,29 | 63,39 | 4018,082 | 149,94 | –86,56 | 7491,963 |
| 1995 | 22 | 562 | 432,03 | –129,97 | 16891,62 | 63,39 | –193,36 | 37386,59 |
| 1995 | 23 | 708,2 | 452,78 | –255,42 | 65241,37 | –129,97 | –125,46 | 15739,24 |
| 1995 | 24 | 327,3 | 473,52 | 146,22 | 21380,28 | –255,42 | 401,64 | 161317,8 |
| сума |  | 5639,17 | 5639,17 | 0 | 220462,4 | –146,22 | 190,50 | 436529,03 |

Розрахуємо d–статистику на основі табличних даних:

.

Користуючись статистичними таблицями Дарбіна–Уотсона для *n*=24 (обсяг вибірки), *k*=1 (кількість незалежних змінних) та заданим рівнем значущості – 5%, знаходимо значення *dl*=1,16 та *du*=1,33. Розрахункове значення *d* знаходиться в інтервалі , тобто (1,33; 4-1,33); (1,33; 2,67). Можна зробити висновок, що автокореляція є відсутньою.

**Завдання для самоперевірки**

1 Наведені дані щодо ДТП в області за 10 місяців року:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| ДТП | 223 | 226 | 241 | 236 | 234 | 243 | 249 | 257 | 252 | 247 |

Побудувати лінійну регресійну модель (лінійний тренд), знайти залишки регресії, коефіцієнт автокореляції залишків. Перевірити модель на наявність автокореляції за критерієм знаків і критерієм Дарбіна-Уотсона з 95% ймовірністю.

2 Зміна курсу долара (приріст, коп.) за тиждень характеризується середніми даними

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| № | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Приріст, коп | 1 | 1,4 | 2 | 1,8 | 1,4 | 0,5 | 0 |

Побудувати лінійну регресійну модель (лінійний тренд), знайти залишки регресії, коефіцієнт автокореляції залишків. Перевірити модель на наявність автокореляції за критерієм знаків і критерієм Дарбіна-Уотсона з 90% ймовірністю.

**Контрольні питання**

1. Для якого типу моделей найбільш характерна властивість автокореляції? Чому?
2. Як визначається вибірковий коефіцієнт автокореляції першого порядку?
3. Опишіть процедуру тестування вибірки на автокореляцію за критерієм знаків.
4. Які параметри використовуються під час перевірки гіпотези щодо існування автокореляції методом Дарбіна-Уотсона?
5. Опишіть процедуру тестування Дарбіна-Уотсона щодо наявності автокореляції в моделі лінійної регресії.