ТЕМА 13. Оцінка якості прогнозу

13.1. Поняття якості прогнозу

13.2. Оцінювання адекватності прогнозованої моделі.

13.3. Критерії визначення якісного прогнозу.

13.4. Оцінка точності прогнозованої моделі та прогнозів.

13.1. Поняття якості прогнозу

Перед тим як використовувати модель для складання реальних прогнозів, її потрібно перевірити на об'єктивність для того щоб забезпечити точність прогнозів. Цю задачу можна вирішити за допомогою наступних дій:

1) визначити, на основі яких статистичних даних побудована модель, потім *фактичні дані порівняти з відповідними оцінками, отриманими за допомогою даної моделі.* Відхилення між фактичними і теоретичними значеннями покажуть, як модель проявить себе при визначенні майбутніх прогнозних значень результуючої змінної. Треба зауважити, що при цьому існує ймовірність того, що висновки можуть бути дещо викривленими, так як модель в середині діапазону даних проявить себе краще, ніж на часових періодах за межами діапазону;

2) результати модельних розрахунків порівняти з фактичними значеннями по мірі надходження інформації. Після отримання нових фактичних даних можна перевірити точність моделі. Недоліком даного підходу є те, що перевірка моделі може зайняти багато часу.

Якість прогнозної моделі є пропорційною точності розрахованих з її допомогою прогнозів. Розрізняють поняття апріорної та апостеорної якості прогнозної моделі.

Апріорну якість оцінюють в умовах відсутності інформації про результати використання моделі.

Апостеорна якість може бути оцінена після практичного використання моделі, тобто розрахунку прогнозу.

Для визначення апріорної якості моделі прогнозують можливі результати її використання за заданих умов при фіксованих значеннях факторних ознак. Тобто визначення якості прогнозної моделі передбачає розрахунок двох прогнозів – прогнозу використання моделі та прогнозу значень ознак за заданих умов.

При цьому якість моделі може бути описана двома характеристиками:

1) *точністю*, тобто збігом фактичних і теоретичних значень показників на етапі побудови моделі;

2) *надійністю*, яка визначається ймовірністю реалізації відповідної оцінки прогнозу. Ймовірність реалізації оцінюється або суб'єктивно (експертні оцінки), або шляхом урахування інтервалу довіри прогнозу.

На практиці при порівнянні декількох прогнозних моделей дослідники, як правило, обмежуються тільки аналізом їх **точності**. Для цього розраховують залишкову варіацію (суму квадратів відхилень обчислених значень від фактичних), середньоквадратичне значення відхилень результатів прогнозу показника від його фактичних значень та довірчі інтервали. Якщо ж модель використовується для

прогнозування значень нових об'єктів, то традиційні апріорні критерії не завжди будуть гарантувати її високу якість. В таких випадках розраховують показники середнього та емпіричного ризику.

Для оцінки апостеорної якості моделі використовують ***параметричні*** та ***непараметричні*** методи.

Оптимальний прогноз — це зроблене на підставі економічної теорії передбачення, яке використовує всю доступну на момент побудови прогнозу інформацію. Для оптимального прогнозу граничний виграш та граничні витрати збігаються.

Оптимальний прогноз іще називають прогнозом раціональних сподівань. Раціональні сподівання можуть відрізнятися від фактичних значень, але будь-яка різниця має бути випадковою й непередбачуваною. Оскільки раціональні сподівання ґрунтуються на коректній економічній теорії, вони мають властивості *незсуненості* (за умови квадратичної функції витрат) та *ефективності*.

Незсуненість означає, що помилка прогнозу має нульове математичне сподівання.

Ефективність передбачає, що в процесі прогнозування буде використана вся доступна інформація, отже, помилка прогнозу не буде корелювати з цією інформацією.

Існують численні критерії перевірки раціональності послідовності прогнозів. Стандартний критерій незсуненості потребує перевірки гіпотези стосовно того, що $\alpha=0$ та $\beta = 1$ водночас для такої моделі:

$$y_t = \alpha + \beta \hat{y}_t + \varepsilon_t \quad (1)$$

де y_t начень або спостережень;

$\hat{y}_t$ - ряд прогнозованих значень;

ε_t — випадкові залишки.

Перевірка ефективності є складнішою оскільки неможливо коректно визначити відповідний масив інформації, стосовно якого похибки прогнозу будуть некорельованими.

Узагальнюючи огляд критеріїв визначення якісного прогнозу, можна стверджувати, що варто користуватися системою критеріїв, які мають враховувати:

- ✓ кількість зусиль, витрачених на побудову моделі, і наявність готових комп'ютерних програм;
- ✓ швидкість, із якою метод уловлює істотні зміни у поведінці ряду, наприклад раптовий зсув математичного сподівання або збільшення кута нахилу лінії тренду;
- ✓ існування серійної кореляції у помилках;
- ✓ незмінюваність первинних даних;
- ✓ повний обсяг роботи в деяких сферах діяльності — тисячі рядів щомісяця потребують оновлення, невеликі витрати й швидкість мають першорядне значення;
- ✓ терміновість прогнозування.

13.2. Оцінювання адекватності прогнозованої моделі

Незалежно від виду і способу побудови економіко-математичної моделі питання про можливість її застосування з метою аналізу і прогнозу економічного явища може бути вирішено тільки після встановлення адекватності моделі. Оскільки повної відповідності моделі реальному процесу або об'єкту бути не може, адекватність певною мірою умовне поняття. При моделюванні мається на увазі адекватність не взагалі, а адекватність тим властивостям моделі, які вважаються суттєвими для дослідження.

Модель згладжування $\hat{y}_t$ певного часового ряду y_t вважається адекватною, якщо правильно відображає систематичні компоненти часового ряду. Ця вимога еквівалентна вимозі, щоб залишкова компонента $e_j = y_j - \hat{y}_j$ (t=1, 2, ..., n) відповідала таким властивостям випадкової компоненти часового ряду як:

1) випадковість коливань рівнів залишкової послідовності,
2) відповідність розподілу випадкової компоненти нормальному закону,
3) рівність математичного сподівання випадкової компоненти нулю,
4) незалежність значень рівнів випадкової компоненти.

Розглянемо, яким чином здійснюється перевірка цих властивостей залишкової послідовності.

1) Перевірка випадковості коливань рівнів залишкової послідовності означає перевірку гіпотези про правильність вибору виду тренду. Для дослідження випадковості відхилень від тренду розраховують набір різниць $e_j = y_j - \hat{y}_j$, (t=1, 2, ..., n). Характер цих відхилень вивчається за допомогою ряду непараметричних критеріїв. Одним з таких критеріїв є «критерій серій», який використовує медіану вибірки, критерій «зростаючих» та «спадних» серій тощо.

Згідно з критерієм серій ряд із величин e_t розташовують у порядку зростання їхніх значень і знаходять медіану e_m одержаного варіаційного ряду, тобто значення, що перебуває в середині для непарного n або середню арифметичну з двох середніх значень для n парного. Повертаючись до вхідної послідовності e_t і порівнюючи значення цієї послідовності з e_m, ставлять знак «плюс», якщо значення e_t , перевищує медіану, і знак «мінус», якщо воно менше за медіану; у випадку однаковості порівнюваних величин відповідне значення e_t пропускають. Отже, одержують послідовність, що складається із плюсів та мінусів, загальна кількість яких не перевищує n. Послідовність розташованих одне за одним плюсів або мінусів називають серією. Щоб послідовність e_t була випадковою вибіркою, довжина найдовшої серії не має бути занадто великою, а загальна кількість серій – занадто малою.

Позначимо довжину найдовшої серії через K_{max}, а загальну кількість серій – через v. Вибірка вважається випадковою, якщо виконуються такі нерівності для 5%-го рівня значущості:

$$K_{\max} < [3{,}3(\ln n + 1)], \qquad (2)$$

$$v > [\frac{1}{2}(n+1-1{,}96\sqrt{n-1})], \quad (3)$$

де квадратні дужки означають цілу частину числа.

Якщо хоча б принаймні одна з цих нерівностей порушується, то гіпотеза про випадковий характер відхилень рівнів часового ряду від тренду спростовується, а модель тренду визнається неадекватною.

2) Перевірка відповідності розподілу випадкової компоненти нормальному закону розподілу може бути зроблена лише наближено за допомогою дослідження показників **асиметрії (*A*) і ексцесу (*E*)**, оскільки часові ряди, як правило, не дуже довгі. При нормальному розподілі показники асиметрії і ексцесу певної генеральної сукупності дорівнюють нулю.

Неперервна випадкова величина має нормальний розподіл, якщо її щільність розподілу дорівнює $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$, де $a,\sigma>0$ - параметри.

Графік функції $f(x)$ називається кривою нормального розподілу.

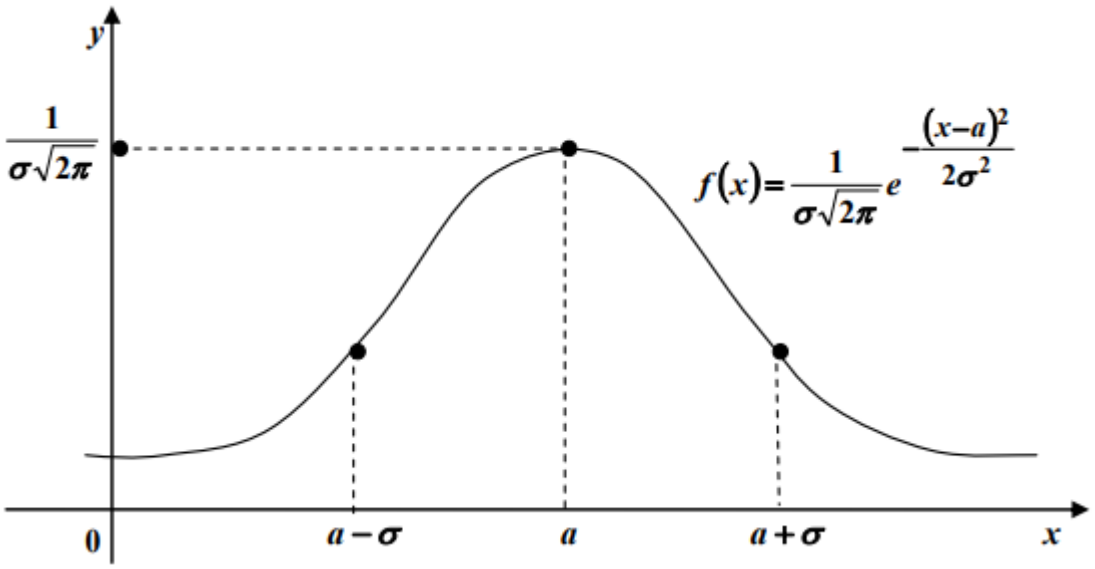


Рис. 1. Графік кривої нормального розподілу

Згідно цього закону основна кількість досліджуваних даних по кожному показнику повинна бути згрупована навколо їх середнього значення; дані з дуже малими та дуже великими значеннями повинні зустрічатися якомога рідше. Для кількісної оцінки ступеня відхилення інформації від нормального розподілу використовують відношення показника асиметрії до її похибки та відношення показника ексцесу до його похибки.

Показник асиметрії A розраховують за формулою:

$$A=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^3}{n\cdot\sigma^3} \quad (4)$$

Похибку асиметрії m_a визначають наступним чином:

$$m_a=\sqrt{\frac{6}{n}} \quad (5)$$

Криві, зображені на рисунку 2, дозволяють ілюструвати симетрію і два найбільш поширених види асиметрії розподілу. При симетричному розподілі (а) середня арифметична, мода і медіана рівні між собою. Для асиметричних кривих ці статистичні величини неоднакові. Причому середня арифметична і медіана зміщені від центра вбік довшої вітки кривої. Оскільки середня арифметична $\bar{x}$ «чутлива» до «точного» положення більш віддалених від моди (M_o) точок кривої, а медіана (M_e) «нечутлива», то середня ($\bar{x}$) зрушена більше, ніж медіана (M_e). У цьому випадку медіана знаходиться між модою і середньою арифметичною.

В статистичному аналізі мода і медіана – це узагальнюючі характеристики сукупності, які відрізняються особливим розташуванням у варіаційному ряду розподілу. Їх ще називають структурні (позиційні) середні.

Модою називають значення ознаки, що має найбільшу частоту в статистичному ряду розподілу.

Медіаною називають таке значення ознаки, яке поділяє проранжований ряд розподілу на дві рівні частини, тобто значення, яке перебуває в середині ряду розподілу.

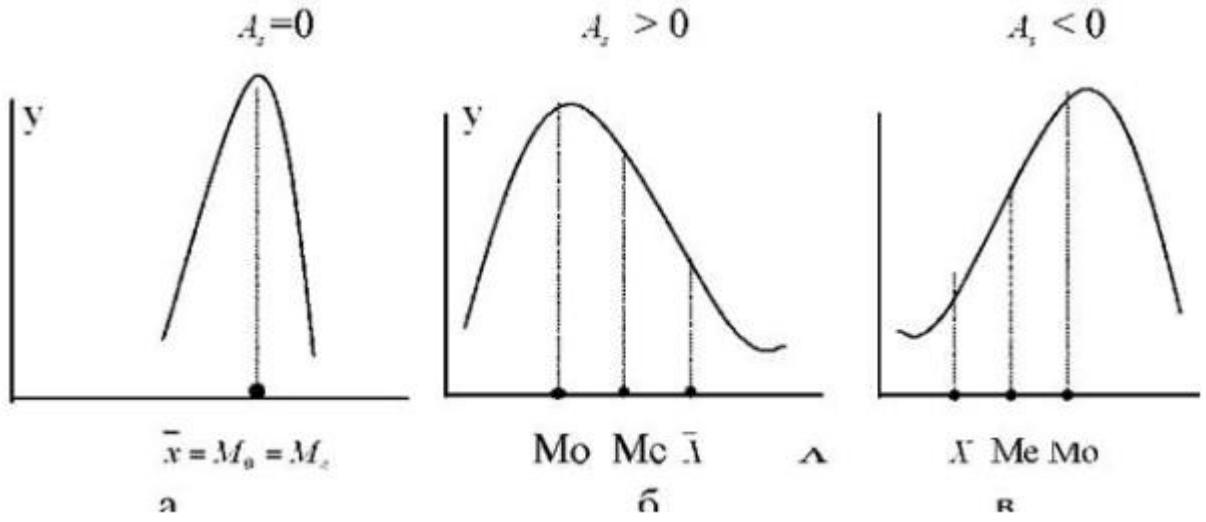


Рис. 2. Форми розподілу при різних значеннях асиметрії (А)

З рис. 2, видно, що коли $A > 0$, ми маємо правобічну асиметрію, а коли $A < 0$ – лівосторонню.

Від'ємна асиметрія свідчить, що у масиві даних показника переважають дані з великими значеннями, а дані з малими значеннями зустрічаються рідше. У додатній асиметрії тенденція зворотна, тобто у масиві значень показника дані з малими значеннями зустрічаються частіше, ніж з великими.

Отже, напрямок асиметрії геометрично встановлюється дуже просто. Кількісна форма ступеня асиметрії вимагає знаходження її алгебраїчної міри.

Якщо асиметрія більше 0,5, то незалежно від знаку вона вважається значною; асиметрія менше 0,25 вважається незначною. Крім того, характер асиметрії вказує на напрямок розвитку. При дослідженні варіації ознак, щодо яких є зацікавленість в їх збільшенні (виконання норм, випуск продукції тощо), правобічна асиметрія свідчить про прогресивність розвитку – про те, що воно йде в бік збільшення показника, а лівостороння асиметрія, отже, вказує на регресивний розвиток. При дослідженні варіації ознак, щодо яких є зацікавленість в їх зменшенні (собівартість,

трудомісткість, витрата сировини на одиницю продукції тощо), правобічна асиметрія свідчить про недоліки в розвитку досліджуваного процесу, лівобічна – про прогресивність його розвитку, про те, що останній йде в бік зменшення показника.

Для встановлення міри відхилення від нормального розподілу вираховують показник ексцесу (*E*). Він характеризує відхилення від нормального розподілу варіантів із виступанням або падінням вершини кривої розподілу.

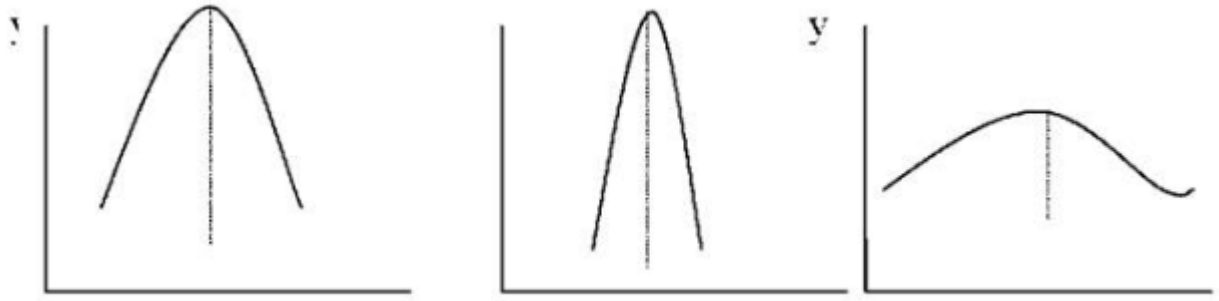
Показник ексцесу *E* розраховують за формулою:

$$E = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{n \cdot \sigma^4} \qquad (6)$$

Похибку ексцесу m_e відповідно за формулами:

$$m_e = \sqrt{\frac{24}{n}} \text{ або } m_e = 2 \cdot m_a \qquad (7)$$

За нормального розподілу показник ексцесу E=0. Якщо $E > 0$, то дані густо згруповані навколо середньої і крива розподілу буде гостро вершинною; і навпаки, якщо $E < 0$, то крива розподілу буде плоско вершинною (рис. 3).



а) $E=0$ б) $E > 0$ в) $E < 0$

Рис. 3. Форми розподілу для різних значень ексцесу

Однак, якщо відношення $\frac{A}{m_a}$ та $\frac{E}{m_e}$ менші 3, то асиметрія і ексцес не мають суттєвого значення і досліджувана інформація підпадає під нормальний закон розподілу.

Припустимо, що відхилення від тренду являють собою вибірку з генеральної сукупності, тому можна визначити тільки вибіркові характеристики асиметрії й ексцесу та їхні помилки:

$\hat{A}$ (вибіркова характеристика асиметрії):

$$\hat{A} = \frac{\frac{1}{n}\sum_{t=1}^{n} e_t^3}{\sqrt{\left(\frac{1}{n}\sum_{t=1}^{n} e_t^2\right)^3}} \qquad (8)$$

$\sigma_{\hat{A}}$ середньоквадратична помилка вибіркової характеристики асиметрії:

$$\sigma_{|\hat{A}|} = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}} \qquad (9)$$

$E\hat{}$ (вибіркова характеристика ексцесу):

$$\hat{E} = \frac{\frac{1}{n}\sum_{t=1}^{n} e_t^4}{\left(\frac{1}{n}\sum_{t=1}^{n} e_t^2\right)^2} - 3 \qquad (10)$$

$\sigma_{\hat{E}}$ помилка вибіркової характеристики ексцесу:

$$\sigma_{\hat{E}} = \sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}}, \qquad (11)$$

Якщо одночасно виконуються такі нерівності:

$$\left|\hat{A}\right| < 1{,}5\sigma_{\hat{A}} \qquad (12)$$

$$\left|\hat{E} + \frac{6}{n+1}\right| < 1{,}5\sigma_{\hat{E}}, \qquad (13)$$

то гіпотеза про нормальний характер розподілу випадкової компоненти приймається.

Якщо виконується хоча б одна з цих нерівностей:

$$\left|\hat{A}\right| \geq 2\sigma_{\hat{A}} \qquad (14)$$

$$\left|\hat{E} + \frac{6}{n+1}\right| \geq 2\sigma_{\hat{E}} \qquad (15)$$

то гіпотеза про нормальний характер розподілу відхиляється, трендова модель визнається неадекватною. Інші випадки потребують додаткової перевірки за допомогою складніших критеріїв.

Для адекватних моделей доцільно ставити запитання щодо оцінювання їхньої точності. Вважається, що моделі з меншою розбіжністю між фактичними й розрахунковими значеннями краще відображають досліджуваний процес у майбутньому.

Для характеристики рівня близькості використовують такі описові статистики:

середнє квадратичне відхилення (або дисперсія)

$$\hat{\sigma}_e = \sqrt{\frac{1}{n-k-1}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \qquad (16)$$

середню відносну похибку апроксимації (чим ближче до 0, тим точніша модель):

$$\overline{e}_{відн.} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| 100\% \qquad (17)$$

коефіцієнт сходження:

$$\varphi^2 = \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \overline{y}_t)^2} \qquad (18)$$

коефіцієнт детермінації (чим ближче до 1, тим точніша модель):

$$R^2 = 1 - \varphi^2 \qquad (19)$$

У формулах (16-19)

n — кількість рівнів ряду,

k — кількість пояснювальних змінних у моделі,

$\hat{y}_t$ — оцінки рівнів ряду за моделлю,

$\overline{y}$ — середнє арифметичне значення вибірки.

На підставі розглянутих показників можна з кількох адекватних моделей обрати найточнішу. Помилка прогнозу, обчисленого для періоду, характеристики якого вже були використані при оцінюванні параметрів моделі, як правило, буде незначною та мало залежатиме від теоретичної обґрунтованості, застосованої для побудови моделі.

Оскільки формально-статистичний вибір кращої моделі в багатьох випадках не гарантує цілковитої впевненості в його правильності, адже добрий прогноз можна отримати і на підставі поганої моделі, і навпаки, тому про якість застосовуваних методик і моделей у прогнозуванні можна судити лише за сукупністю зіставлень прогнозів і їх реалізації. ***При цьому незалежно від обраної методики та моделі прогнозування джерелами помилок прогнозу можуть бути:***

1) природа змінних (випадковий характер змінних гарантує, що прогноз відхилятиметься від справжніх величин, навіть якщо модель правильно специфікована, її параметри точно відомі);

2) природа моделі (сам процес оцінювання спричиняє похибки оцінок параметрів);

3) помилки, привнесені прогнозом незалежних випадкових величин (пояснювальних змінних);

4) помилки специфікації моделі.

3) Перевірка рівності математичного сподівання випадкової компоненти нулю, якщо вона розподілена за нормальним законом, виконується на основі *t-критерію* Стьюдента. Розрахункове значення цього критерію задається формулою

$$t = \frac{\bar{e} - 0}{\sigma_e}\sqrt{n} \qquad (20)$$

де $\bar{e}$ – середнє арифметичне значення рівнів залишкової послідовності e_t;

σ_e – стандартне (середньоквадратичне) відхилення для цієї послідовності.

Якщо розрахункове значення t менше табличного значення t_α статистики Стьюдента із заданим рівнем значущості α і числом ступенів свободи $n–1$, то гіпотеза про рівність нулю математичного сподівання випадкової послідовності приймається, в протилежному випадку ця гіпотеза відхиляється і модель вважається неадекватною.

4) Перевірка незалежності значень рівнів випадкової компоненти, тобто перевірка відсутності суттєвої автокореляції в залишковій послідовності, може здійснюватися за рядом критеріїв, найбільш поширеним з яких є DW-критерій Дарбіна-Уотсона. Зазначимо, що розрахункове значення критерію Дарбіна-Уотсона в інтервалі від 2 до 4 свідчать про від'ємний зв'язок. У цьому випадку його треба перетворити за формулою DW' = 4 – DW і далі використовувати значення DW'.

Висновок про адекватність трендової моделі робиться, якщо всі розглянуті вище чотири перевірки властивостей залишкової послідовності дають позитивний результат.

13.3. Критерії визначення якісного прогнозу

Якість прогнозу характеризують такі поширені в прогностичній літературі терміни, як точність і надійність. Проте зміст цих термінів часто тлумачать досить неоднозначно. Це можна пояснити тим, що нині поки не знайдено ефективного підходу до оцінювання якості прогнозу, окрім його практичного підтвердження.

Про точність прогнозу прийнято судити за розміром помилки прогнозу – різниці між прогнозним і фактичним значенням досліджуваного показника. Але такий підхід можливий лише тоді, якщо дослідник має інформацію стосовно справжніх значень часового ряду, який він оцінював під час розроблення прогнозів. Наприклад, період випередження вже завершився, і дослідник має фактичні значення змінної (це можливо в разі короткотермінового прогнозування) або прогноз перебуває в стадії розроблення, тобто прогнозування здійснюється для певного моменту часу в минулому, для якого існують фактичні дані. Спрощену схему періодів прогнозування показано на рис. 4.

В останньому випадку йдеться про використання *expost*-прогнозу. Його сутність полягає у побудові моделі за меншим обсягом даних (n - m) із подальшим порівнянням прогнозованих оцінок за останніми m точками (для t від n - m + *1* до n) із відомими фактичними, але спеціально залишеними рівнями ряду. Отримані ретроспективно помилки прогнозу певною мірою характеризують точність застосовуваної методики прогнозування й можуть виявитися корисними в зіставленні кількох прогнозів.

Рис. 4. Спрощена схема періодів прогнозування

Розмір помилки ретроспективного прогнозу не можна розглядати як остаточний доказ придатності, або навпаки, непридатності застосовуваного методу прогнозування. До неї варто ставитися з певною обережністю і при її застосуванні в якості міри точності необхідно враховувати, що вона отримана при використанні лише частини наявних даних. Проте ця міра точності має більшу наочність і теоретично більш надійна, ніж похибка прогнозу, обчислена для періоду характеристики котрого вже були використані при оцінюванні параметрів моделі.

На основі розглянутих показників можна зробити вибір з декількох адекватних моделей найбільш точної. Помилка прогнозу, обчислена для періоду, характеристики котрого вже були використані при оцінюванні параметрів моделі, як правило, будуть незначними і мало залежатимуть від теоретичної обґрунтованості, застосованої для побудови моделі.

13.4. Оцінка точності прогнозної моделі та прогнозів

Щодо оцінювання помилки прогнозу слід вважати, що оцінюється одна змінна *y*. Існує певна вибірка вимірювань даної змінної розміром N. Дана вибірка формується на певному відрізку часу – періоді підстави прогнозу. Тому з кожним номером спостереження пов'язаний момент часу. Отже, розмір N, крім визначення кількості ретроспективних спостережень, також визначає довжину часового відрізка.

Основні характеристики, які використовуються для оцінювання якості прогнозу в зазначеній ситуації, доцільно розглянути більш детально.

Першу групу утворюють ***абсолютні показники помилки прогнозу***. Вони дозволяють кількісно визначити величину розбіжності між прогнозом і фактом спостереження в одиницях вимірюваної змінної. Розрізняють такі характеристики:

1) Помилка (похибка) прогнозу

$$e_j = y_j - \hat{y}_j \qquad (21)$$

де y_j – фактичне значення змінної *j*-того виміру; $\hat{y}_j$ – прогнозне значення досліджуваної змінної.

2) Абсолютна помилка (похибка) прогнозу

$$\Delta_j = \left| y_j - \hat{y}_j \right| \qquad (22)$$

3) Середня абсолютна похибка прогнозу (the mean absolute error):

$$MAE = \frac{\sum_{j=1}^{N} \left| y_j - \hat{y}_j \right|}{N} = \frac{1}{N} \sum_{j=1}^{N} \Delta_j \qquad (23)$$

4) Сума квадратів помилки (sum square error):

$$SSE = \sum_{j=1}^{N} (y_j - \hat{y}_j)^2 = \sum_{j=1}^{N} e_j{}^2 \qquad (24)$$

5) Середній квадрат помилки (похибки) (mean square error):

$$MSE = \frac{1}{N} \sum_{j=1}^{N} (y_j - \hat{y}_j)^2 = \frac{1}{N} \sum_{j=1}^{N} e_j{}^2 \qquad (25)$$

Найчастіше два останні показники (*SSE* та *MSE)* використовуються при виборі оптимального методу прогнозування. У більшості статистичних пакетів дані показники використовуються як критерії вибору моделі.

6) Середнє квадратичне відхилення похибки (the root mean square error):

$$RMSE = \sqrt{\frac{\sum_{j=1}^{N} (y_j - \hat{y}_j)^2}{N}} = \sqrt{\frac{1}{N} \sum_{j=1}^{N} e_j{}^2} \qquad (26)$$

Значення всіх перерахованих вище показників залежать від масштабу вимірювань, який у ряді випадків (зокрема, при міжоб'єктних зіставленнях) зменшує об'єктивність оцінок. Для того щоб уникнути цього, використовують відносні показники вимірювань помилки прогнозу, виражені в частках одиниці або у відсотках. Дані показники утворюють другу групу – ***відносні помилки прогнозу:***

1) Відносна помилка прогнозу:

$$\varepsilon_j = \frac{\left| y_j - \hat{y}_j \right|}{y_j} \times 100 = \frac{\Delta_j}{y_j} \times 100 \qquad (27)$$

2) Середня абсолютна похибка у відсотках або відсоткова помилка (mean absolute percentage error):

$$MAPE = \frac{1}{N} \sum_{j=1}^{N} \frac{\left| y_j - \hat{y}_j \right|}{y_j} \times 100 \qquad (28)$$

Показник, як правило, використовується для порівняння точності прогнозів різнорідних об'єктів прогнозування, оскільки в ньому застосовуються відносні величини. Вважається, що прогноз має високу точність, якщо *MAPE* < 10 %. Прогноз має гарну точність, якщо значення даного показника знаходиться між 10 і 20 %. Прогноз має задовільну точність за умови, що 20 % > *MAPE* < *5*0 %. Якщо значення показника більше за 50%, то такий прогноз має незадовільну точність.

3) Корінь з середньоквадратичної похибки у відсотках (the root mean square percentage error):

$$RMSPE = \sqrt{\frac{1}{N} \sum_{j=1}^{N} \left(\frac{y_j - \hat{y}_j}{y_j} \right)^2} \times 100 \qquad (29)$$

Чим менше значення цих величин, тим вища якість ретро-прогнозу. На практиці ці характеристики використовують досить часто. Даний підхід дає гарні результати, якщо на періоді ретро-прогнозу не виникають принципово нові закономірності. На підставі двох критеріїв MAPE, RMSPE можна дійти висновку стосовно загального рівня адекватності моделі шляхом їх порівняння. Цей рівень наведений у таблиці 1.

4) Середня відсоткова помилка (похибка) прогнозу (mean percentage error):

$$MPE = \frac{1}{N}\sum_{j=1}^{N}\frac{y_j - \hat{y}_j}{y_j}100\,\% \qquad (30)$$

MPE характеризує відносний ступінь зміщеності прогнозу. За умови, що втрати при прогнозуванні, пов'язані із завищенням фактичного майбутнього значення, врівноважуються заниженням, ідеальний прогноз повинен бути незміщеним, і обидві міри повинні прагнути до нуля. Середня відсоткова помилка не визначена при нульових даних і не повинна перевищувати 5 %

Таблиця 1

Точність прогнозу в залежності від MAPE, RMSPE

MAPE, RMSPE	Точність прогнозу
Менше 10%	Висока
10% - 20%	Добра
20% - 40%	Задовільна
40% - 50%	Погана
Більше 50%	Незадовільна

Вадою обговорених вище характеристик точності прогнозів є їх залежність від обраних одиниць виміру. Було б корисним указати безрозмірний показник, аналогічний до коефіцієнта кореляції. Одним з таких показників є коефіцієнт невідповідності Тейла, чисельником якого є середньоквадратична похибка прогнозу, а знаменник дорівнює квадратному кореню із середнього квадрата фактичних та оцінних значень. Таким чином, ***третю групу утворюють коефіцієнти невідповідності (розбіжності) запропоновані А. Тейлом.***

$$K_T = \sqrt{\frac{\sum_{j=1}^{N}(y_j - \hat{y}_j)^2}{\sum_{j=1}^{N} y_j^2}} \qquad (31)$$

або

$$K_T = \sqrt{\frac{\sum_{j=1}^{N}(y_j - \hat{y}_j)^2}{\sum_{j=1}^{N} y_j^2 + \sum_{j=1}^{N} \hat{y}_j^2}} \qquad (32)$$

Друга форма коефіцієнта Тейла (32) є більш поширеною на практиці. Якщо не зроблено спеціальних застережень, то за замовчуванням використовується саме вона. Слід додати, що на практиці в ряді випадків дослідження об'єктів прогнозування й їх динаміки більш ефективними методами оцінювання точності прогнозів є обчислення

коефіцієнтів невідповідності Тейла в приростах досліджуваного показника y_i. Неважко переконатися, що «зроблений прогноз» має коефіцієнт Тейла, що дорівнює 0.

Перевага коефіцієнта Тейла полягає в тому, що його значення завжди перебувають у межах від нуля до одиниці. Якщо всі прогнози абсолютно точні, то він дорівнює 0. Якщо всі прогнози дорівнюють нулю, а жодне з фактичних значень не дорівнює нулю або навпаки, коефіцієнт дорівнюватиме одиниці. Таким чином, мале значення коефіцієнта засвідчує, що прогноз є точним, але максимального значення не існує. Значення, яке дорівнює одиниці, відповідає ситуації, коли всі прогнозові значення дорівнюють нулю, що нереально під час прогнозування номінальних величин, але під час розгляду змін такий прогноз відповідає моделі «без змін». Більші за одиницю значення вказують на те, що прогноз гірший, ніж прогноз «без змін».

Коефіцієнт невідповідності Тейла (K_T) може бути розкладений на три частини: пропорцію зсунення

$$U^M = \frac{(y_i - \hat{y}_i)^2}{\frac{1}{m}\sum_{i=1}^{m}\frac{(y_i - \hat{y}_i)}{y_i}} \qquad (33)$$

пропорцію дисперсії

$$U^S = \frac{(\sigma_y - \sigma_{\hat{y}})^2}{\frac{1}{m}\sum_{i=1}^{m}\frac{(y_i - \hat{y}_i)}{y_i}}, \qquad (34)$$

пропорцію коваріації

$$U^C = \frac{2(1-\rho)(\sigma_y \cdot \sigma_{\hat{y}})}{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}. \qquad (35)$$

Зазначимо, що $U^M + U^S + U^C = 1$. Критерій зсуву пропорції (U^M) використовується, щоб перевірити, чи є систематичне відхилення середніх розрахованих та фактичних рядів, тобто чи дає модель систематично завищені або занижені прогнози. Чим менше значення, тим краще. Якщо U^M дорівнює нулю, у розрахованих (прогнозних) значеннях немає зсунень, тобто з моделлю все гаразд. Пропорція дисперсії (U^S) використовується, щоб переконатися, що модель має достатні динамічні властивості для відтворення дисперсії фактичних рядів. Наприклад, модель може відтворювати систематично менші коливання, ніж фактичні. Як і у випадку критерію U^M, менше значення U^S вказує на менше зсунення. Пропорція коваріації вказує, як корелюють фактичні та розраховані ряди. Якщо U^C дорівнює 1, то фактичні та розраховані ряди корелюють ідеально.

Критичні точки важливі як критерії якості, оскільки деякі моделі можуть бути точними, але погано передбачати зміни тенденції (наприклад, поворотні точки в циклах), тобто погано відтворювати критичні точки. Інші моделі можуть бути неточними, але мати гарний динамічний характер. Загалом може бути певний компроміс між точністю та динамічними властивостями моделі.

Формального тесту для оцінки цієї властивості не існує. Проте візуальний огляд розрахованих та фактичних рядів звичайно одразу виявляє, здатна модель відтворювати критичні точки чи ні.

Четверта група характеристик пов'язана з порівнянням окремих значень досліджуваної змінної з її середнім значенням. До таких показників відносять:

абсолютне відхилення від середньої:

$$AD_j = \left| y_j - \bar{y} \right| \tag{36}$$

середнє абсолютне відхилення (*mean absolute deviation):*

$$MAD = \frac{1}{N} \sum_{j=1}^{N} \left| y_j - \bar{y} \right|. \tag{37}$$

Всі зазначені показники відносяться до *параметричних методів аналізу точності прогнозів.* Обговорені характеристики точності прогнозів є параметричними в тому сенсі, що вони потребують виконання заданих припущень щодо властивостей математичного сподівання та дисперсії, чинних за умов нормальності відповідних розподілів. Наприклад, використовуючи *MSE*, ми неявно припускаємо, що всі похибки прогнозу мають однакові й постійні математичні сподівання та дисперсії. У реальних економічних ситуаціях найчастіше порушуються припущення гомоскедастичності та відсутності автокореляції. Можна стверджувати, що кожного разу прогноз будується у новій ситуації, отже, порівняння числової точності прогнозів, зроблених у різні моменти часу, не зовсім коректне. Наведені міркування зумовили використання непараметричних методів аналізу точності прогнозів.

Непараметричні методи аналізу точності прогнозів.

Непараметричні методи не залежать від вигляду розподілу, тож не потребують припущення щодо нормальності розподілів. Це особливо корисно, коли йдеться про дані, які унеможливлюють використання числових шкал. Розглянемо два типи непараметричних критеріїв: критерій знаків та рангові критерії.

Критерій знаків для порівняння точності двох послідовностей прогнозів базується на відсотку випадків, коли метод визначення прогнозу *A* кращий, ніж метод *B*. Таке порівняння здійснюють для індивідуальних прогнозів однакових подій (змінних). Якщо обидва методи дають однакову точність, імовірність відповіді «так» на запитання «чи прогноз А кращий за прогноз *B*» становить 0,5 для кожного з *m* випадків прогнозування. Число *K* випадків, коли прогноз *A* кращий, підпорядковано біноміальному розподілу ймовірностей

$$p(K = x) = C_m^x 0{,}5^x 0{,}5^{m-x}. \tag{38}$$

Отже, можна підрахувати імовірність того, що $K \geq x$. Якщо довжина послідовності прогнозів значна, для оцінювання ймовірностей можна використати нормальну апроксимацію біноміального розподілу.

Критерій знаків можна також використовувати для перевірки значущості описової статистики, відомої під назвою «відсоток кращих результатів», яка показує

відсоток випадків, у яких один метод прогнозування кращий за інший і розраховується за формулою:

$$\eta = \frac{m}{m+n} \tag{39}$$

де m — кількість прогнозів, підтверджених фактичними даними;

n — кількість прогнозів, не підтверджених фактичними даними.

Коли всі прогнози підтверджуються, $n=0$ і $\eta=1$; якщо всі прогнози не підтвердилися, то m, а отже, і η дорівнюють 0.

Рангові критерії. У разі застосування цих критеріїв чисельна характеристика точності (абсолютна похибка, коли маємо один прогноз, або *MSE*, коли розглядають послідовність прогнозів) замінюється рангами, які потім перевіряють на значущість. Наприклад, якщо послідовності прогнозів показників *A* та *B* одержують за допомогою *k* методів, то спочатку обчислюють *MSE*, потім їхні значення ранжують від *1* (найменша *MSE*) до *k* (найбільша *MSE*) (відповідні ранги позначають через R_A, та R_B, для $i = 1, \ldots, k$). Після знаходження різниць (d_i) між рангами обчислюють *коефіцієнт рангової кореляції Спірмена:*

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \tag{40}$$

За нульову гіпотезу приймають відсутність залежності між рангами, тобто жоден з методів не є гіршим за решту. Гіпотеза відкидається, якщо значення r_s досить велике.

Хоча непараметричні методи мають свої переваги, важливо усвідомлювати, що вони ігнорують частину доступної інформації. Так, критерії знаків та рангів не враховують числових значень похибок.