УДК 338.27:004(045)

МОДЕЛІ ПРОГНОЗУВАННЯ ФІНАНСОВИХ ПРОЦЕСІВ НА ОСНОВІ РЕГРЕСІЙНИХ МОДЕЛЕЙ ТА НОВІТНІХ КОМП'ЮТЕРНИХ ТЕХНОЛОГІЙ

Лисогор В.М., д.т.н., ***Фаюра Н.Д.***, к.е.н., ***Голишева А.В.***
Вінницький державний аграрний університет

Описано однопараметрична прогресійна модель функцій відкугу, що характеризує об'єм фінансових послуг за даний період часу. Визначення достовірності отриманих оцінок, що містяться в регресійній моделі дає змогу оцінити ступінь їх точності. В цілому, розроблений математичний апарат є ефективним засобом аналізу даних з метою фінансового прогнозування.

Характерною рисою сучасного етапу становлення суспільства є інформатизація. Зростаючі вимоги у представленні економічної інформації потребують застосування статистичних методів для перевірки висунутих гіпотез та побудови математичних моделей різних явищ та процесів.

Так, в [1,2] створена модель колекційного аналізу показників діяльності комерційних структур: фондових бірж, середніх та малих підприємств, телекомпаній та ін. В [3,4] наведений математичний інструментарій для вирішення різноманітних задач з використанням статичних методів, розроблені приклади аналізу економічних даних з допомогою Microsoft Excel. Проте актуальною залишається задача прогнозування в економічних дослідженнях за допомогою сучасних методів аналізу даних.

Мета дослідження – обґрунтування та розробка одно параметричної лінійної регресивної моделі для прогнозування поведінки економічної системи в наступний період з використанням методів статичного аналізу та інформаційних технологій.

Основний результат – отримання математичної моделі, яка найбільш точно описує поведінку економічної системи під впливом певного фактору за допомогою методів регресивного аналізу.

Згідно [3], слід виділити декілька важливих моментів, які необхідно враховувати при побудові залежностей і подальшому економічному прогнозуванні на їх основі:

– сукупність факторів, які діють на характеристику, що досліджується, не повинна змінюватись на протязі прогнозованого періоду (тобто ухвалення урядом законів, які змінюють умови економічної діяльності, економічний спад можуть зробити прогнозування некоректним);

– із збільшенням періоду передбачення поведінки економічної системи знижується точність прогнозування.

Для прогнозування поведінки економічної системи на протязі певного періоду будемо використовувати одно параметричну регресійну модель.

Основною метою регресивного аналізу є отримання функції залежності вихідної величини (відгуку) від зміни певного фактору, яким у більшості випадків є час. Так, наприклад, якщо ми здійснили комп'ютерний моніторинг щомісячного об'єму реалізації певного виду товару за останній рік і отримали n вибіркових даних спостереження : (X_1, Y_1), (X_2, Y_2),…,(X_n, Y_n), які зосереджені поблизу деякої лінії , тоді функція буде мати вигляд:

$$\widehat{Y} = A_0 + A_1X \quad (1)$$

де $\widehat{Y}$ – величина, передбачуваної регресивної моделі, X – значення керованого фактора.

Необхідно отримати такі значення коефіцієнтів A_0 та A_1, при яких сума квадратів похибок є мінімальною. Похибкою в даному випадку є відстань по вертикалі від даної точки до отриманої лінійної залежності. Опишемо значення функції відгуку конкретних точках:

$$\widehat{Y}1 = A0 + A1X1;\ \widehat{Y}2 = A0 + A1X2;...;\ \widehat{Y}n = A0 + A1Xn \quad (2)$$

У цьому випадку можна сформувати похибку лінійної моделі у такому виді:

$$\widetilde{y}1 = \widehat{Y} - Y1 = A0 + A1X1 - Y1;\ \widetilde{y}2 = \widehat{Y}2 - Y2 = A0 + A1X1 - Y2;....; \quad (3)$$

Сформуємо квадратичну функцію невязки нашої системи у вигляді:

$$J = \widetilde{y}_1^2 + \widetilde{y}_2^2 + ... + \widetilde{y}_n^2 \quad (4)$$

Підставивши конкретні значення з виразу (3) в (4) отримаємо:

$$J = (A_0 + A_1X_1 - Y_1)^2 + (A_0 + A_1X_2 - Y_2)^2 + ... + (A_0 + A_1X_n - Y_n)^2 \quad (5)$$

В загальному вигляді квадратична функція похибки буде мати вигляд:

$$J = \sum_{i=1}^{n}(A_0 + A_1X_i - Y_i)^2 \quad (6)$$

Для отримання невідомих значень параметрів A_0 та A_1, при яких квадратична функція J є мінімальною, застосуємо звичайні методи математичного аналізу. Умовою мінімуму є:

$$\partial J / \partial A_0 = 0; \partial J / \partial A_1 = 0 \quad (7)$$

Підставивши конкретні значення (6) у вираз (7) отримаємо:

$$\frac{\partial J}{\partial A_0} = \frac{\partial}{\partial A_0}\sum_{i=1}^{n}(A_0 + A_1X_i - Y_i)^2 = \sum_{i=1}^{n}\frac{\partial}{\partial A_0}(A_0 + A_1X_i - \text{Y}i)^2 =$$

$$\sum_{i=1}^{n}2(A_0 + A_1X_i - Y_i) = 2(nA_0 + A_1\sum_{i=1}^{n}X_i - \sum_{i-1}^{n}Y_i) = 0 \quad (8)$$

отже, кінцевий результат буде мати вигляд:

$$nA_0 + (\sum_{i=1}^{n}X_1)A_1 = \sum_{i=1}^{n}Y_1 \quad (9)$$

Відносно значення другого невідомого параметра A_1 діємо аналогічно.

$$\frac{\partial F}{\partial A_1} = \sum_{i=1}^{n}\frac{\partial}{\partial A_1}(A_0 + A_1X_i - Y_i)^2 = \sum_{ш=1}^{m}2X_i(A_0 + A_1X_i - Y_i)$$

$$= 2A_0\sum_{i=1}^{n}X_i + 2A_1\sum_{i=1}^{n}X_i^2 - 2\sum_{i=1}^{n}Y_iX_i = 0 \quad (10)$$

звідки

$$(\sum_{i=1}^{n} X_i)A_0 + (\sum_{i=1}^{n} X_1^2) = \sum_{i=1}^{n} X_i\, Y_i \tag{11}$$

Звівши результати виразів (9) та (11), отримаємо лінійну регресійну модель для визначення невідомих параметрів A_0 та A_1. В матричному вигляді ці рівняння будуть мати вигляд:

$$\left\| \begin{matrix} n & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & \sum_{i=1}^{n} X_1^2 \end{matrix} \right\| \cdot \left\| \begin{matrix} A_0 \\ A_1 \end{matrix} \right\| = \left\| \begin{matrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_i Y_i \end{matrix} \right\| \tag{12}$$

Вирішуючи рівняння (12), отримуємо:

$$A_0 = \frac{\sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_i Y_i}{n\sum_{i=1}^{n} X_1 - (\sum_{i=1}^{n} X_1)^2} \tag{13}$$

$$A_1 = \frac{n\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_1)^2} \tag{14}$$

де n – кількість вибірок експериментальних даних.

Звичайно, мірою помилки регресивної моделі є стандартне (середньоквадратичне) відхилення s, яке обчислюється за формулою:

$$s = \sqrt{\sum_{i=1}^{n} E_1^2 /(n-2)} = \sqrt{\sum_{i=1}^{n} (A_0 + A_1 X_i - Y_i)^2 /(n-2)}. \tag{15}$$

Для нормального розподілення процесів, згідно [5], приблизно 66% точок знаходиться в межах одного стандартного відхилення від моделі і 95% точок – в рамках двох стандартних відхилень.

Визначимо коефіцієнти функцій відгуку, яка характеризує об’єм реалізації товару в задані періоди часу. В таблиці 1 наведені вихідні та розрахункові дані.

Таблиця 1

Період надання послуг, місяці	Об’єм фінансових послуг, тис., шт..	Розрахункові дані	
X_i	Y_i	X_i^2	X_iY_i
1	9,69	1	6,7
2	8,76	4	13,8
3	7,28	9	21,84
4	8,86	16	32,8
5	8,31	25	41,55
6	6,58	36	44,4
7	7,72	49	54,04
8	7,28	64	58,24
9	9,8	81	77,4
10	9,03	100	89
11	9,2	121	101,2
12	10,27	144	114
Сума 78	95,99	650	654,99

Процес знаходження коефіцієнтів функції відгуку згідно наведеного вище математичного апарату є досить трудомістким. Тому, для скорочення часу і спрощення процесу розрахунки доцільно застосувати можливість аналізу даних програми Microsoft Excel. Так, за допомогою функцій СУММ, СТЕПЕНЬ КОРЕНЬ ми швидко отримали розрахункові дані, представлені в таблиці 1 та визначили згідно формул (13-14) A_0 =6,59 та A_1 = 0,22.

Далі, на основі отриманих коефіцієнтів, побудували математичну модель об'єму реалізацій товарів на протязі певного період, що описується рівнянням У=6,59+0,22Х.

Для знаходження середньоквадратичного відхилення s згідно формули (15) можна по черзі ви користати функції СУММ та КОРЕНЬ, або скористатись однією функцією СТАНДОТКЛОН. В даному випадку, середня квадратична похибка s склала 0,51. Це означає, що 66% точок знаходиться в межах $\Delta\hat{Y} \pm 0{,}51$, а 95% точок в межах $\Delta\hat{Y} \pm 1{,}02$.

Наступним кроком після розробки регресивної моделі є аналіз достовірності моделі. Іншими словами, необхідно встановити чи є гіпотеза покладена в основу регресивної моделі, достатньо вірною, щоб з її допомогою зробити передбачення і який ступінь точності можуть мати оцінки, що містяться в регресивній моделі

Аналіз точності та достовірності регресивних моделей, згідно [4], можна провести шляхом розгляду наступних характеристик: розкид даних навколо їх середнього значення $SS_1 = \sum_{i=1}^{n}(Y_1 - \bar{Y})^2$

– розкид даних навколо лінії регресії:

$$SS_r = \sum_{i=1}^{n}(Y_1 - \hat{Y}_1)^2$$

– розкид лінії регресії навколо середнього значення даних:

$$SS_b = \sum_{i=1}^{n}(\hat{Y}_1 - \bar{Y})^2$$

Для цих видів розкиду справедливим є відношення виду:

$$SS_1 \approx SS_r + SS_b$$

Отже, з допомогою вбудованих функцій Microsoft Excel СУММ, СРЗНАЧ, СТЕПЕНЬ ми знайшли SS_i =9,37, SS_r=2,64, SS_b=6,74. При цьому була виконана рівність $SS_1 \approx SS_r + SS_b$

Методика перевірки достовірності моделі була розроблена Фішером. Регресивна модель може розглядатися як гіпотеза, заснована на деякій сукупності експериментальних даних. Методика Фішера включає правило зіставлення даної гіпотези з вибіркою експериментальних даних, що дозволяє встановити непридатність гіпотези у випадку її малої вірогідності.

Введене позначення $F_{y,b,a}$ квантиль розподілу Фішера з Y і з B степенями свободи при визначенні ймовірності помилки a. Величини F містяться в спеціальних статистичних таблицях.

В даному випадку Y=1, b=n-k-1, де n – число точок експериментальних даних; k – кількість змінних керувань; наприклад, у випадку регресивної моделі, з'єднуючої змінні X і У, k=1, тоді в нашому випадку б=12-1-1=10; а – визначена ймовірність помилки – 95%. Отже, на основі цих даних знайдемо із таблиці статистичного розподілу значення $F_{y,b,a=}$ 4,96.

Далі застосовуючи методику Фішера перевіряємо умову $MS_b/MS_r > F_{y,b,a}$

В нашому випадку 36,04>4,96. Отже, умова виконується, що свідчить про достовірність моделі описаної функції відгуку.

Таким чином, розроблена одно параметрична регресійна модель дозволяє встановити вид функції відгуку, що характеризує об'єм продажу певного виду товару за визначений період часу. Визначення достовірності отриманих оцінок, що містяться в регресійній моделі дає змогу оцінити ступінь її достовірності. Отже, розроблений математичний апарат є ефективним засобом аналізу даних з метою економічного прогнозування.

Література:

1.Практическая бизнес-статистика: Пер. с англ. – М.: Издательский дом «Вильямс», 2002. – 1056с.

2.Сигел Флейс Дж. Статистические методы для изучения таблиц, долей и пропорций – М: Фининсы и статистика, 1989. – 319с.

3.Поллард Дж. Справ очник по вычислительным методам статистики. – М.: Финансы и статистика, 1982. – 344с.

4.Лаппач С.Н., Чубенко А.В., Бабич П.Н. Статистика в науке и бизнесе . – К.:Морион, 2002. – 640с.

5.Барковський В.В., Барковська Н.В., Лопатин О.К. Територія ймовірностей та математична статистика. – К.: ЦУЛ,2002. – 448с.

6.Лисогор В.Н., Яремко С.А., Прогнозування процесів із застосуванням методів статистичного аналізу та новітніх комп'ютерних технологій.// Збірник наукових праць Всеукраїнської науково-практичної конференції «Пріоритети економічного розвитку України: історія та сьогодення 10-11 квітня 2007 . – Вінниця, 2007 – С.503-509.