УДК 330.43

КОРЕЛЯЦІЙНО-РЕГРЕСІЙНИЙ АНАЛІЗ ОБСЯГІВ ЗБУТУ ПРОДУКЦІЇ ПРОМИСЛОВОГО ПІДПРИЄМСТВА

Харченко Ю.А., к.т.н.
E-mail: yuranchartsch@mail.ru
Полтавський національний технічний університет імені Юрія Кондратюка

У статті досліджуються напрями підвищення ефективності та конкурентоспроможності промислових підприємств в сучасних умовах. Обґрунтовано необхідність застосування множинного кореляційно-регресійного аналізу для кількісної оцінки взаємозалежностей між статистичними ознаками щоб збільшити точність прогнозу та зменшити мінливість показника. Розроблена економіко-математична модель прогнозування і планування обсягів збуту окремих видів продукції з метою удосконалення системи формування виробничих планів на майбутній короткотерміновий період та збільшення доходу підприємства. Наведено приклад розрахунку для одного виду продукції. Виконана перевірка на наявність лінійної залежності між усіма або кількома факторними ознаками (мультиколінеарність). Здійснено апробацію моделі при удосконаленні системи планування та збуту продукції одного з машинобудівних підприємств м. Полтави. Запропоновано напрями її використання для моделювання різних варіантів управління збутом продукції та поліпшення фінансового стану

Ключові слова: економіко-математична модель, кореляційно-регресійний аналіз, мультиколінеарність, управління, прогнозування, планування

UDC 330.43

CORRELATION REGRESSION ANALISIS OF THE VOLUME OF SALES OF INDUSTRIAL ENTERPRISE

Kharchenko Y.A., PhD in Engineering
E-mail: yuranchartsch@mail.ru
Poltava National Technical Yuri Kondratyuk University

The areas of improving the efficiency and competitiveness of industrial enterprises in modern conditions have been considered in the article. Necessity of using multiple correlation-regression analysis for quantitative mark the interdependencies between statistical characteristics has been substantiated for increase accuracy of prediction and reducing variability of index. The economic-mathematical model of prediction and planning of sales of certain products aiming at the improvement system of formation production plans for future short-term period and increase income of the enterprise has been developed. The example of calculation for one type of product has been given. Checking for linear dependence between all or several factor variable (multicollinearity) has been made. The

© Харченко Ю.А., к.т.н., 2014

approbation of the model at improvement of system planning and sales of a machine-building enterprise of Poltava has been carried out. The areas its use for modeling different variants of administration sales of products and improve the financial situation have been proposed

Keywords: economic-mathematical model, correlation-regression analysis, multicollinearity, administration, prediction, planning

Актуальність проблеми. Сучасний етап розвитку ринкової економіки в Україні потребує використання науково-обґрунтованих управлінських рішень, які підвищують ефективність та конкурентоспроможність промислових підприємств. Оцінити вплив різноманітних чинників на кінцевий результат можна за допомогою економіко-статистичних методів. Їх застосування дає можливість провести аналіз досліджуваних статистичних показників та на основі отриманих даних побудувати математичну модель.

Велика кількість діючих факторів обумовлює необхідність застосування множинного кореляційно-регресійного аналізу для кількісної оцінки взаємо залежностей між статистичними ознаками, що характеризують окремі соціально-економічні процеси. Під час аналізу необхідно встановити теоретичну форму зв'язку між факторними і результативними ознаками (регресійний аналіз) та визначити тісноту цього зв'язку (кореляційний аналіз), тобто кількісно виміряти й оцінити механізм взаємодії факторних ознак. Параметри кореляційного аналізу використовується для цілеспрямованого регулювання рівнів результативних ознак.

Таким чином, отримані результати кореляційно-регресійного аналізу можна використати для обґрунтування управлінських рішень щодо прогнозування і планування обсягів реалізації окремих видів продукції в короткотерміновому майбутньому періоді.

Аналіз останніх наукових досліджень. Проблемі підвищення точності прогнозів за даними кореляційно-регресійного аналізу присвячена велика кількість наукових праць. Основні статистичні поняття, методи лінійного та багатофакторного регресійного аналізу, а також деякі напрямки практичного застосування цих методів при ухваленні управлінських рішень наведено Д.Е. Ханком в роботі [1]. Методи прогнозування соціально-економічних процесів та приклади практичного застосування математичних моделей для розв'язання економічних задач представлені Хемді А. Таха в праці [2].

Основні принципи моделювання та моделі прогнозування, які можна застосувати для розв'язання управлінських задач, наведені Д. Муром в [3]. Фундаментальні основи регресійного аналізу, а також методи побудови та дослідження моделей різного ступеня складності з аналізом практичних аспектів їх застосування подані Н. Дрейпером в роботі [4].

В роботах П.В Ґудзя та Р.І. Заворотнього [5; 6] використані регресійні факторні моделі для удосконалення систем управління підприємствами вітчизняної економіки. У праці Д.С. Терехова [7] досліджуються особливості застосування інформаційних систем при оптимізації діяльності промислових підприємств. Методи економіко-математичного моделювання використані в роботах А.С. Пятигіна та В.Г. М'ячина [8; 9] для оцінювання середовища прийняття управлінських рішень та при прогнозуванні попиту на інноваційну продукцію машинобудівного підприємства.

Актуальність подальших досліджень полягає в розробленні економіко-математичної моделі прогнозування і планування обсягів збуту окремих видів продукції для підвищення рівня обґрунтованості управлінських рішень.

Мета роботи: розробити економіко-математичну модель прогнозування і планування обсягів збуту окремих видів продукції з метою удосконалення системи формування виробничих планів на майбутній короткотерміновий період та збільшення доходу промислового підприємства.

Викладення основного матеріалу дослідження. Завданням кореляційно-регресійного аналізу є побудова та аналіз економіко-математичної моделі рівняння регресії (рівняння кореляційного зв'язку), що відображає залежність результативної ознаки від кількох факторних ознак і дає оцінку міри щільності зв'язку.

Кореляційні зв'язки встановлюються в середньому для великої сукупності даних з інформаційної бази, яка має достатньо типові та надійні статистичні характеристики, а також якісну однорідність (наближеність умов формування результативних і факторних ознак) та кількісну однорідність (відсутність одиниці спостереження, яка за числовими характеристиками суттєво відрізняється від основної маси даних). Ці особливості потребують розв'язання двох задач: знаходження форми функціонального зв'язку та визначення міри наближення кореляційного зв'язку за ним.

Однією з найбільш простих і розповсюджених моделей є лінійна регресія, але вона, як правило, не може забезпечити необхідну точність прогнозування. Щоб збільшити точність прогнозу та зменшити мінливість показника використовуються багатофакторні моделі (множинна кореляція). При виборі незалежних змінних (факторів) в таких моделях потрібно враховувати наявність зв'язку з залежною змінною (показником) та відсутність тісного зв'язку з будь-якою іншою незалежною змінною, тобто взаємної кореляції. Фактори повинні відображати різні аспекти досліджуваного процесу. Для аналізу щільності зв'язку в багатофакторній кореляційно-регресійній моделі складають матрицю парних коефіцієнтів кореляції, які вимірюють щільність лінійного зв'язку кожного фактора з результативною ознакою і з кожною з решти ознак-факторів (кореляційна матриця). За формою зв'язку розрізняють кореляційні зв'язки прямі й обернені, лінійні й нелінійні, одно й багатофакторні. Прямі й обернені зв'язки розрізняють залежно від напряму зміни результативної ознаки при зміні факторної ознаки. Якщо співпадають напрями – прямий зв'язок, якщо ні – обернений. Залежно від характеру зміни показника Y при зміні фактора X виділяють лінійні та нелінійні зв'язки.

Сукупність інформаційних вхідних даних необхідно перевірити на наявність лінійної залежності між усіма або кількома факторними ознаками (мультиколінеарність). В економічних процесах об'єктивно існують співвідношення між окремими факторами. Мультиколінеарність, як правило, проявляється в стохастичній (прихованій) формі. Її наявність призводить до серйозного зниження точності оцінок параметрів регресії, скривлення оцінки дисперсії залишків, дисперсії коефіцієнтів регресії і коваріації між ними.

Коефіцієнти регресії стають ненадійними, їх неможливо трактувати як міру впливу відповідного фактора на незалежну змінну. Оцінки стають дуже чутливими до вибіркових даних, тобто невелике збільшення об'єму вибірки може спричинити до значних змін в значеннях оцінок.

Отже потрібно виконати відповідну перевірку, тому що отримана регресія буде характеризуватися великою мінливістю та певною надлишковістю (розраховані значення більше очікуваних). Під час перевірки на мультиколінеарність спочатку статистичні значення факторних ознак X_i замінюються на стандартизовані (нормалізовані) дані. Потім виконується побудова кореляційної матриці, яка складається з коефіцієнтів кореляції, що обчислюються для кожної можливої пари

змінних. Аналіз отриманих залежностей дає можливість оцінити ступінь та напрямок взаємозв'язку між факторами. Далі обчислюється визначник кореляційної матриці, який вказує на можливу кореляцію між факторами (якщо значення близьке до нуля).

Загальна мультиколінеарність в масиві виявляється за критерієм Пірсона (χ^2) з надійністю $p = 0{,}95$. Потім визначаються частинні коефіцієнти кореляції, які характеризують тісноту зв'язку між двома змінними за умови, що третя не впливає на цей зв'язок. Далі за допомогою t-статистики з надійністю $p = 0{,}95$ виявляються пари факторів, між якими існує мультиколінеарність.

Після встановлення стохастичної мультиколінеарності по можливості її потрібно усунути. Одним з таких методів є метод вилучення змінних (факторів). Його суть полягає в видаленні (повинно узгоджуватися з метою дослідження та економічною доцільністю) однієї або кількох високо-корельованих пояснюючих змінних з регресії. Потім нова модель заново оцінюється. Для перевірки виконується побудова кількох можливих моделей та обчислюється значення коефіцієнта детермінації R^2, що вимірює частку варіації показника, яка пояснюється взаємозв'язком між незалежною змінною і факторами та розрахункове значення критерію Фішера F_p за формулами:

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{Y}_i - \overline{Y})^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}, \qquad (1)$$

де $\hat{Y}_i$ – розрахункове значення показника для і-го спостереження;

Y_i – статистичне значення показника для і-го спостереження;

$\overline{Y}$ – вибіркове середнє значення показника.

$$F_p = \frac{R^2}{1 - R^2}\left(\frac{n - m - 1}{m}\right), \qquad (2)$$

де n – кількість спостережень;

m – кількість факторів.

Чим ближче значення R^2 до одиниці, тим краще статистичні дані відповідають побудованій функції регресії. Значення F_p повинно бути в

кілька разів більше від табличного, тоді регресія буде не тільки значимою, а й мати практичну цінність для прогнозування.

Після видалення однієї змінної будується двофакторна модель. Коефіцієнти рівняння регресії $Y = a_0 + a_1 X_1 + a_2 X_2$ обчислюються з використанням матричних операцій за формулою:

$$\vec{a} = \left[[X]^T [X] \right]^{-1} [X]^T \vec{Y} \tag{3}$$

Після знаходження оцінок параметрів регресії a_0, a_1, a_2 виконується перевірка нової моделі на мультиколінеарність. Якщо результат позитивний виконуються перевірки адекватності моделі (рівняння регресії) за критерієм Фішера та значущості параметрів (окремих факторів) регресії за критерієм Стьюдента.

Потім обчислюється прогнозне значення показника та інтервальний прогноз (надійний інтервал) для заданих значень факторів за формулами (4 і 6):

$$DY = t_c \, S \sqrt{X_p [X^T X]^{-1} X_P^T + 1}, \tag{4}$$

де t_c – табличне значення критерію Стьюдента;

X_P – вектор-рядок з прогнозними значеннями факторів;

S – стандартне відхилення оцінки (вимірює розсіювання значення показника відносно лінії функції регресії) є квадратним коренем з дисперсії адекватності:

$$S = \sqrt{\frac{\sum_{i=1}^{n} (Y - \hat{Y}_i)^2}{n - m - 1}}, \tag{5}$$

$$Y_{\max(\min)} = Y_P \pm DY, \tag{6}$$

де Y_p – точковий прогноз показника.

Розраховані частинні коефіцієнти регресії показують на середню зміну залежної змінної для одиничної зміни відповідної незалежної змінної та постійних значеннях інших незалежних змінних. За відомим рівнянням регресії можна моделювати майбутні параметри показника при зміні значень одного із факторів. Далі для прогнозу обчислюються частинні коефіцієнти еластичності.

З метою врахування динаміки зміни обсягів реалізації окремих видів продукції при складанні оперативних виробничих планів на

короткотерміновий майбутній період на основі кореляційно-регресійного аналізу була побудована модель. В ній використовуються дані обсягів реалізації продукції за 12 місяців 2013 року одного з машинобудівних підприємств м. Полтави, які мають досить велику мінливість. Дослідження статистичних даних показника Y (обсягу збуту продукції) виявили певні коливання фактичної реалізації окремих видів продукції, що пов'язано з проблемами збуту в одну із країн СНД, на яку припадає найбільша частка експорту. Після первинного оброблення даних з інформаційної бази в якості факторів прийняті: середня ціна продукції (X_1), фонд заробітної плати на виготовлення продукції (X_2), витрати на збут продукції (X_3). Дані з інформаційної бази про реалізацію одного з основних видів продукції наведені в табл. 1.

Таблиця 1. *Дані про реалізацію продукції П1*

Період	**Дохід Y, грн.**	**Фактори X, грн.**		
		Середня ціна	Фонд заробітної плати	Витрати на збут продукції
1	828183,0	330,61	79021,78	24341,20
2	1030589,0	314,11	115901,70	29983,79
3	822400,0	325,25	89684,11	25997,09
4	1036455,0	304,67	99442,89	26242,49
5	840820,0	358,36	84759,81	35088,69
6	866697,0	367,80	88760,86	43634,62
7	1095990,3	306,95	103492,43	47483,37
8	1227065,0	346,92	134108,01	40729,56
9	974255,5	341,54	109763,70	43804,87
10	1041248,0	355,55	121792,57	44906,99
11	2115316,0	265,74	218887,29	52933,20
12	311604,2	450,90	42183,20	9622,36

Результати перевірки мультиколінеарності в масиві факторів за критерієм Пірсона представлені на рис. 1.

ВИЗНАЧНИК КОРЕЛЯЦІЙНОЇ МАТРИЦІ			
R=	0,21763		
ЗНАЧЕННЯ КРИТЕРІЮ XI-КВАДРАТ			
XI2p=	**13,979**	XI2табл=	**7,815**

Рис. 1*. Результати перевірки мультиколінеарності*

Розрахункове значення XI^2 більше табличного, тому в масиві незалежних змінних існує мультиколінеарність в сукупності. Потім за критерієм Стьюдента виконується перевірка мультиколінеарності пари факторів, отримані дані подані на рис. 2.

ЗНАЧЕННЯ КРИТЕРІЮ СТЬЮДЕНТА			
t12=	**-2,053**		
t13=	**-0,142**		tтабл
t23=	**2,425**		**2,262**

Рис. 2. *Перевірка мультиколінеарності пари факторів*

Розрахункове значення t_{23} перевищує табличне значення Стьюдента, тобто існує мультиколінеарність пари факторів. Одним із способів усунення мультиколінеарності є видалення одного фактора з моделі. Для подальшого аналізу вибирається двофакторна модель $Y = f(X_1, X_2)$, яка має кращі значення коефіцієнта детермінації R^2=0,98 (близьке до одиниці) і розрахункового значення критерію Фішера F_p=233,6 (набагато перевищує табличне значення).

Отримана модель знову перевіряється на наявність мультиколінеарності. Розрахункове значення XI^2=7,12 не перевищує табличного значення. Потім обчислюються параметри регресії за формулою (3), які наведені на рис. 3.

a0=	**674963,43**
a1=	**-1633,12**
a2=	**8,34**

Рис. 3. *Параметри регресії моделі*

Далі розраховуються значення t-статистики, вони всі перевищують табличне значення, а отже є значимими (рис. 4).

t0=	**2,71**
t1=	**2,82**
t2=	**13,49**
tтабл=	**2,262**

Рис. 4. *Розраховані значення t-статистики*

За отриманими даними будується прогноз для значень факторів X_1=350,00 грн. і X_2=105,00 грн., який складає Y_p=978664,40 грн. і знаходиться в межах довірчого інтервалу (рис. 5).

Розраховані коефіцієнти еластичності k_1=-0,584 і k_2=0,894 вказують на те, як зміна одного фактора на 1% впливає на значення показника при незмінному значенні іншого фактора. При перевірці результатів прогнозування на інших статистичних даних генеральної сукупності (за 2014 рік) з інформаційної бази похибка прогнозу не перевищує значень середньоквадратичних похибок, які використовувалися при визначенні функції регресії.

Головною метою проведення кореляційно-регресійного аналізу є визначення можливих варіантів управління збутом продукції, а також оцінка можливих шляхів досягнення потрібного результату. Розроблена модель може бути використана для поліпшення планування потреб в матеріальних ресурсах, а також удосконалення процесів підготовки виробництва, якщо буде відомий прогноз збуту продукції на майбутній період. Прогноз повинен мати динамічний характер та адаптуватися до змін з врахуванням останніх даних.

Рис. 5. *Прогноз показника* Y_p

Таким чином більш точне планування обсягів реалізації продукції надасть можливість зменшити витрати. Використання результатів моделювання та прогнозування обсягів збуту продукції підвищить ефективність управлінських рішень та зменшить ймовірність помилкових рішень.

Висновки. Отже побудовану економетричну (регресійну) модель можна використовувати для прогнозування майбутнього обсягу збуту окремого виду продукції та формування відповідної виробничої програми на майбутній короткотерміновий період. Запропонована модель повинна бути інтегрована в існуючу систему підтримки прийняття рішень.

Крім цього кореляційно-регресійний аналіз дає можливість оцінити існуючий стан за допомогою рівняння регресії. Використовуючи дані про величину і напрямок дії факторів, що аналізуються, можна отримати засоби для оцінки та відповідного коригування поточної стратегії планування обсягів випуску всієї номенклатури продукції.

Таким чином кореляційно-регресійний аналіз надає потужний та гнучкий інструмент для вивчення взаємозв'язків між показником та множиною незалежних змінних. Основна мета використання цього методу – краще зрозуміти теперішній стан, відповідно навчитися управляти подіями, що відбуваються, а також точніше передбачити майбутнє.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Ханк Д.Э. Бизнес-прогнозирование, 7-е издание. / Д. Э. Ханк, Д.У. Уичерн, А.Дж. Райтс; перевод с англ. − М.: Издательский дом «Вильямс», 2003. − 656 с.
2. Таха Хемди А. Введение в исследование операций, 7-е издание. / Хемди А. Таха; перевод с англ. − М.: Издательский дом «Вильямс», 2005. − 912 с.
3. Мур Д. Экономическое моделирование в Microsoft Excel, 6-е издание. / Д. Мур, Д. Р. Уэдерфорд, Г. Эллен, Ф. Гулд, Ч. Шмидт; перевод с англ. − М.: Издательский дом «Вильямс», 2004. − 1024 с.
4. Дрейпер Н. Прикладной регрессионный анализ, 3-е издание. / Н. Дрейпер, Г. Смит; перевод с англ. − М.: Издательский дом «Вильямс», 2007. − 912 с.
5. Ґудзь П.В. Аналіз бюджетного управління на підприємстві хімічної промисловості / П.В. Ґудзь, Р.П. Науменко // Бізнесінформ. − 2013. − №12. − С. 144-149.
6. Заворотній Р.І. Запровадження факторних моделей в оцінюванні зростання вітчизняного бізнесу / Р.І. Заворотній // Актуальні проблеми економіки. − 2012. − №3. − С. 312-321.
7. Терехов Д.С. Удосконалення інформаційного забезпечення управління машинобудівним підприємством на основі процесного підходу / Д.С. Терехов // Науковий журнал Економічний часопис-XXI. − 2012. − №11-12. − С. 48-52.
8. Пятигін А.С. Використання аналітичних моделей для оцінювання середовища прийняття управлінських рішень у сфері фінансової діяльності підприємств / А.С. Пятигін // Актуальні проблеми економіки. − 2011. − №11. − С. 308-321.
9. М'ячин В.Г. Прогнозування попиту на інноваційну продукцію машинобудівного підприємства / В.Г. М'ячин, А.Г. Жукова // Економічний простір: Збірник наукових праць. − № 68. − Дніпропетровськ: ПДАБА, 2012. − С. 215-224.