

УДК 519.233.4:65.011.47:001.12:33
© 2010

Опря А.Т., доктор економічних наук
Полтавська державна аграрна академія

МЕТОДОЛОГІЧНІ ОСОБЛИВОСТІ ВИКОРИСТАННЯ ДИСПЕРСІЙНОГО МЕТОДУ В АНАЛІЗІ Й ДОСЛІДЖЕННІ ЕКОНОМІЧНИХ ЯВИЩ І ПРОЦЕСІВ: МОЖЛИВОСТІ Й ОБМЕЖЕННЯ

Рецензент – доктор економічних наук, професор В.Я. Плаксієнко

Метод дисперсійного аналізу, як і інші методи математичної статистики, має особливість (специфіку) у своєму використанні при вирішенні аналітичних завдань у галузі економіки. Тут він може виконувати як основні, так і допоміжні аналітичні функції. А саме: кількісне вимірювання впливу факторних ознак; визначення вірогідності впливу та його довірчих меж у причинно-наслідкових моделях; аналіз окремих середніх і статистична оцінка їх різниці; забезпечення науково обгрунтованого підходу при застосуванні кількісних методів статистики в аналізі й дослідженні економічних явищ і процесів.

Ключові слова: метод дисперсійного аналізу, галузь економіки, кількісне вимірювання впливу факторних ознак, вірогідність впливу, багатофакторний аналіз, кореляційно-регресійний метод, ймовірно-статистичні методи.

Постановка проблеми. В епоху бурхливого розвитку економіки використання методів математичної статистики в аграрно-економічних дослідженнях стає нагальною необхідністю. Треба визнати, що останнім часом широкого застосування у багатофакторному аналізі набув кореляційно-регресійний метод. Водночас майже зовсім не використовується досить ефективний спосіб статистико-математичної обробки даних дослідження — дисперсійний аналіз. Як і інші ймовірно-статистичні методи, він набагато розширює можливості економістів в аналізі виробництва й значно підвищує рівень наукових досліджень.

Основне призначення дисперсійного аналізу — статистично виявити вплив різних факторів на мінливість ознаки, що вивчається. Особливий інтерес становить використання цього методу в аналізі економічних процесів і явищ, коли мінливість результативної ознаки зумовлена одночасною дією кількох факторів із неоднаковою силою впливу. Це, зокрема, спостерігається при аналізі результативних синтетичних показників економічної ефективності виробництва. Найбільш ефективний тут одночасний дисперсійний

аналіз усіх відібраних факторів — багатофакторний аналіз. Можна, звичайно, застосувати й попарне порівняння факторів, при якому всі інші ігноруються, однак такий підхід до розв'язання питання не дає змоги виявити існуючу в дійсності множинність ефектів взаємодії.

Використання економістами дисперсійного методу дає змогу розв'язувати досить важливі завдання, виходячи із сучасних вимог до рівня економічного аналізу. У сфері економічних досліджень цей ефективний статистико-математичний засіб повинен зайняти одне з провідних місць насамперед тому, що використання дисперсійного методу може мати самостійне значення. Зокрема, за його допомогою розв'язуються такі завдання: 1) кількісне вимірювання сили впливу факторних ознак та їх сполучень на результативну; 2) визначення вірогідності впливу та його довірчих меж; 3) аналіз окремих середніх та статистична оцінка їх різниці.

У поглибленому економічному аналізі дисперсійний метод може виконувати також і допоміжні функції. У цьому плані його використання відкриває широкі можливості щодо науково обгрунтованого підходу до застосування інших статистичних методів кількісного аналізу.

Як й інші математико-статистичні методи, дисперсійний аналіз являє собою чисто технічний засіб наукового пізнання. І використання його при вивченні економічних процесів передбачає передусім знання суті процесів, розуміння причинно-наслідкових зв'язків між явищами, що вивчаються, а також вміння виділити найбільш важливі сторони взаємозв'язаних і взаємозумовлених економічних процесів і явищ.

Слід відзначити, що в економічних дослідженнях дисперсійний метод ще не набув такого широкого використання, як у біології, зоотехнії, медицині чи техніці, хоча можливості його застосування в сфері економіки досить широкі. Основне призначення дисперсійного аналізу — статистично виявити вплив факторів на варіацію досліджуваних ознак. Особливий інтерес стано-

вить використання цього методу в тих випадках, коли зміна даної ознаки зумовлена одночасно дією факторів, частка впливу яких неоднакова.

У дисперсійному аналізі використовується властивість суми квадратів центральних відхилень. Суть її полягає в тому, що коли кілька повністю незалежних факторів діють одночасно й зумовлюють загальну змінюваність ознаки, то сума окремих дисперсій, що вимірюють їх вплив, дорівнює загальній дисперсії.

Дисперсійний аналіз — це математико-статистичний метод вивчення результатів спостереження, що залежать від різноманітних одночасно діючих факторів. Він створений у 20-х роках ХХ ст. зусиллями Р. Фішера й суттєвий розвиток одержав у працях Іейтса. Основи дисперсійного аналізу у 1933 році були описані М.Ф. Деревиським у додаткових розділах до підручника В. Іогансена «Елементи точного вчення про мінливість та спадковість».

Аналіз основних досліджень і публікацій, у яких започатковано розв'язання проблеми. У дослідженні окремих аспектів статистичної методології в системі аграрної економіки спостерігається два напрями, наукові ідеї й завдання яких взаємопов'язані в частині аналітичних функцій статистики взагалі й дисперсійного методу, зокрема. Для першого напрямку характерна сконцентрованість зусиль науковців на дослідженні теоретико-методологічних питань виміру економічної ефективності виробництва. В аграрному секторі ці питання висвітлені в роботах О.А. Бугуцького, В.Ф. Горянського, А.М. Єріної, М.Я. Кушвіда, А.Т. Опрі, В.П. Трофімова.

Об'єктивними передумовами формування другого напрямку слід вважати потреби у вивченні та оцінці закономірностей розвитку економіки й прогнозування виробництва на підставі вивчення ролі окремих факторів та їх взаємодій у справі підвищення його ефективності. Для даного напрямку досліджень притаманні загальнотеоретичний і прикладний підходи при розробці питань застосування математичних і математико-статистичних методів у вирішенні складних господарських завдань. Найвагоміший внесок тут зробили вчені: А.Я. Боярський, Л.В. Кантарович, В.С. Немчинов, В.В. Новожилов, Й.С. Пасхавер та ін. Окремі питання прикладного характеру розроблені в дослідженнях О.П. Крастина, В.І. Кобринського, А.А. Френкеля й ін. Більшість названих вище дослідників використовувала допоміжні функції дисперсійного методу як критерію надійності результатів досліджень. Щодо використання основних фун-

кцій методу, тобто дослідження та кількісний вимір причинно-наслідкових зв'язків на базі теоретичних засад дисперсійного методу, тут пріоритет належить вченим першого напрямку.

Результати дослідження. Викладене вище не вичерпує можливостей дисперсійного методу аналізу. Знання його особливостей дозволяє безпосередньо оцінити вірогідність тих чи інших розрахунків при використанні методів статистичних групувань, кореляції, регресії, середніх. Особливо широкі можливості дисперсійного аналізу при оцінці множинних кореляційно-регресійних залежностей. Маючи порівняно невелику кількість одиниць спостереження, можна вводити в аналіз ряд ознак-факторів, обчислюючи випадкову помилку на достатньо великому числі ступенів вільності.

Зіставляючи кореляційно-регресійні моделі з двома і більше змінними на невеликій сукупності об'єктів, за допомогою дисперсійного аналізу можна вирішити два досить важливих питання: по-перше, в якому взаємозв'язку знаходяться включені у модель фактори, і, по-друге, чи будуть істотними висновки, зроблені на невеликій вибірці змінних. Ігнорування цього положення забирає чимало часу на пошуки істотних факторів-аргументів, а іноді навіть знецінює економічні дослідження.

Слід зауважити, що метод дисперсійного аналізу має переваги над іншими статистико-математичними методами. Назвемо головні з них. Передусім, використовуючи даний метод у багатофакторному аналізі економічних явищ, можна одержати картину, яка показує вплив кожного фактора в різних умовах, створюваних змінами різних факторів. При цьому застосування різних комбінацій факторів, що вивчаються, дає надійнішу підставу для практичних рекомендацій, які придатні й за мінливих умов. В аналізі економічних явищ, де фактори іноді перебувають у складному переплетінні, дисперсійний метод дає змогу об'єктивно оцінити складні явища, що виникають при такій взаємодії. Наприклад, у трифакторному комплексі мають місце подвійні і потрійні взаємодії факторів.

Розглянемо можливості та обмеження при застосуванні дисперсійного методу в економічному аналізі на прикладі показників урожайності пшениці озимої.

Наведена матриця (див. рис.) схематично зображує окремі середні показники врожайності пшениці озимої (101 підприємство Полісся України), одержані при різних варіантах сполучення факторів: рівня удобреності ґрунту, якості

земель і фондооснащеності досліджуваних підприємств. Так, різниця за рівнем урожайності, зумовлена підвищенням норм внесення мінеральних добрив, кращою якістю ґрунту та вищим рівнем фондооснащеності, становить 10,6 ц/га ($Y_{A_2B_2C_2} - Y_{A_1B_1C_1}$). У господарствах з однаково високими рівнями удобреності та якості землі, але з різним рівнем фондооснащеності абсолютна зміна урожайності пшениці дорівнює 1,4 ц/га ($Y_{A_2B_2C_2} - Y_{A_2B_2C_1}$).

Матриця окремих середніх дає повну інформацію про зміни рівня врожайності при можливих поєднаннях досліджуваних факторів.

Кінцеві статистичні характеристики дисперсійного аналізу, які несуть у собі інформацію про варіацію показників урожайності пшениці озимої та ступінь впливу факторів на її рівень, а також істотність такого впливу, наведено у таблиці 1.

A ₁	A ₁	A ₂	B ₁	B ₂	C ₁	C ₂
A	18,0	18,3	14,2	21,7	16,3	19,6
A ₂		18,6	17,1	20,1	15,5	21,6
B ₁			15,7		15,9	15,3
B			18,3		18,3	
B ₂				20,9		25,8
C ₁					12,3	
C					18,3	
C ₂						20,7
$Y_{A_1B_1C_1} = 15,2$ $Y_{A_1B_2C_1} = 17,5$ $Y_{A_2B_1C_1} = 16,7$ $Y_{A_2B_2C_1} = 14,4$ $Y_{A_1B_1C_2} = 13,3$ $Y_{A_1B_2C_2} = 25,9$ $Y_{A_2B_1C_2} = 17,4$ $Y_{A_2B_2C_2} = 25,8$						

Рис. Матриця окремих середніх показників урожайності пшениці при різних можливих поєднаннях факторів у дисперсійному комплексі:

A – кількість внесених на 1 га ріллі мінеральних добрив, ц діючої речовини; B – оцінка землі, бали; C – вартість основних виробничих фондів на 1 га ріллі, тис. грн.

1. Результати дисперсійного аналізу трифакторного комплексу залежності врожайності пшениці озимої від кількості внесених мінеральних добрив (А), якості землі (В) і фондооснащеності підприємств (С)

Статистичні характеристики	Умовні позначення	Вплив факторів									
		врахований			їх поєднання				сумарний		
		А	В	С	АВ	АС	ВС	АВС	врахований (x)	неврахований (z)	разом (y)
Дисперсія :											
невиправлена	С'	26,6	662,1	532,9	99,7	29,4	641,2	15,2	-	-	-
виправлена	С	26,4	656,8	528,7	98,9	29,2	636,1	15,1	1991	2145	4136
Показник ступеня впливу факторів та їх поєднань	η^2	0,006	0,159	0,128	0,024	0,007	0,154	0,004	0,481	0,519	1
Число ступенів вільності	ν	1	1	1	1	1	1	1	7	87	94
Девіата	σ^2	26,4	656,8	528,1	98,9	29,2	636,1	15,1	284,4	24,6	-
Коефіцієнт Фішера											
розрахунковий	F_p	1,1	26,6	21,4	4,1	1,2	25,8	0,6	11,5	-	-
табличний	F_T	$\begin{Bmatrix} 11,5 \\ 6,9 \\ 3,9 \end{Bmatrix}$	$\begin{Bmatrix} 11,5 \\ 6,9 \\ 3,9 \end{Bmatrix}$	$\begin{Bmatrix} 11,5 \\ 6,9 \\ 3,9 \end{Bmatrix}$	$\begin{Bmatrix} 11,5 \\ 6,9 \\ 3,9 \end{Bmatrix}$	$\begin{Bmatrix} 11,5 \\ 6,9 \\ 3,9 \end{Bmatrix}$	$\begin{Bmatrix} 11,5 \\ 6,9 \\ 3,9 \end{Bmatrix}$	$\begin{Bmatrix} 11,5 \\ 6,9 \\ 3,9 \end{Bmatrix}$	$\begin{Bmatrix} 8,9 \\ 2,8 \\ 2,1 \end{Bmatrix}$	-	-

Вірогідним у нашому випадку виявився вплив якості землі й фондооснащеності підприємств. Перший із названих факторів визначає 15,9% коливання рівня урожайності пшениці, другий – 12,8%. Понад 15% варіації врожайності пояснюється взаємодією цих двох факторів. Вплив поєднання факторів удобреності та якості ґрунту виявився незначним і вірогідним лише при порозі ймовірності 0,95. Одержаної інформації про коливання врожайності (будь-якої культури) цілком достатньо для того, щоб мати підставу продовжити аналіз – залежно від поставленого завдання – у потрібному напрямі його поглиблення.

Як було зазначено вище, природа методу відкриває можливості об'єктивної оцінки застосовуваних в аналізі групувань, визначення істотності коефіцієнтів кореляції і різниці середніх, а також для дослідження множинних кореляційних залежностей, зокрема для оцінки лінійної регресії. Таку оцінку можна виконати, якщо розмежувати загальну суму квадратів відхилень $\sum (y - \bar{y})^2$ на суму квадратів відхилень від пря-

мої лінії теоретичної регресії $\sum (\bar{y}_x - \bar{y})^2$, суму квадратів залишкових відхилень від цієї теоретичної лінії регресії $\sum (y - \bar{y}_x)^2$, а потім із останньої виділити суму квадратів відхилень середніх, які утворюють емпіричну лінію регресії.

Якщо регресія дійсно прямолінійна, то відхилення від неї слід вважати випадковими. У такому разі буде випадковою і частина відхилень, яка припадає на відмінність між теоретичною та емпіричною лініями регресії. Якщо ж теоретична лінія у вигляді прямої неправильно відображує форму зв'язку, то відхилення емпіричної лінії регресії від неї повинні розглядатися не як випадкові, а як закономірне відображення кривизни регресії. Порівняння одержаних відхилень із чисто випадковою їх величиною й дає відповідь на питання про прямолінійність регресії.

Вихідні передумови теорії дисперсійного аналізу зумовлюють також деякі обмеження у його використанні при аналізі економічних показників. Розглянемо їх.

2. Порівняння статистичних характеристик кількісного впливу факторів удобреності ґрунту (А), якості землі (В), фондооснащеності підприємств (С) на врожайність пшениці озимої та оцінку істотності впливу при різних способах розподілу одиниць спостережень у дисперсійному комплексі

Види факторів	Умовні позначення	Вид дисперсійного комплексу									
		рівномірний					нерівномірний				
		η^2	F_p	F_T при P			η^2	F_p	F_T при P		
0,999	0,95			0,99	0,999	0,95			0,99		
Врахованих	А	0,006	<u>0,07</u>	11,5	6,9	3,9	0,002	<u>0,004</u>	11,5	6,9	3,9
	В	0,159	26,65	11,5	6,9	3,9	0,233	35,04	11,5	6,9	3,9
	С	0,128	21,45	11,5	6,9	3,9	0,059	<u>8,90</u>	11,5	6,9	3,9
Поєднань факторів	АВ	0,024	<u>40,9</u>	11,5	6,9	3,9	0,017	<u>2,54</u>	11,5	6,9	3,9
	АС	0,007	<u>1,18</u>	11,5	6,9	3,9	0,004	<u>0,53</u>	11,5	6,9	3,9
	ВС	0,154	25,81	11,5	6,9	3,9	0,0097	<u>10,12</u>	11,5	6,9	3,9
	АВС	0,004	<u>0,61</u>	11,5	6,9	3,9	0,002	<u>0,29</u>	11,5	6,9	3,9
Сумарний :											
врахованих	х	0,481	11,54	3,9	2,8	2,1	0,382	8,21	3,9	2,8	2,1
неврахованих	z	0,519	-	-	-	-	0,618	-	-	-	-
Разом	у	1,000	-	-	-	-	1,000	-	-	-	-

Примітка. У таблиці неістотні показники обведено. Однією лінією підкреслено показники вірогідності при порозі ймовірності 0,95, двома – при порозі ймовірності 0,95 і 0,99.

Істотним недоліком методу є те, що на результати аналізу впливає рівень показників груп, сформованих за досліджуваним фактором. Дисперсійні комплекси, побудовані при одних рівнях факторних градацій, можуть відображувати достовірний вплив, а при інших рівнях такий вплив може бути відсутнім. При цьому виявляється і різниця в показниках ступеня впливу факторів на результативну ознаку. Таким чином, від того, як згруповані дані досліджень у статистичному комплексі, залежать й деякі його результативні характеристики.

Вплив змін рівнів факторних градацій на кінцеві результати розглянемо на прикладі побудови двох видів дисперсійних комплексів – рівномірного і нерівномірного (табл. 2). Такий підхід пояснюється тим, що різні принципи розподілу частот, природно, зумовлюють різні рівні факторних градацій. Вплив зміни рівня останніх на результативні статистичні характеристики дисперсійного комплексу вивчався на прикладі залежності врожайності пшениці озимої від трьох уже розглянутих вище факторів.

Розрахунки показали: якщо дисперсійний комплекс побудований за принципом рівномірного, показник ступеню загального впливу факторів, що аналізуються, за своєю абсолютною величиною на 10% перевищує відповідний показник комплексу, побудованого за принципом нерівномірного розподілу. У нашому прикладі

виявилось чотири (у першому випадку – три) розрахункових показники F-критерію, що були нижчими від своїх табличних значень, тобто при всіх порогах ймовірності, і два – при рівні ймовірності 0,999. Спостерігалася різниця і в показниках ступеню індивідуального впливу досліджуваних факторів та їх поєднань.

Отже, потрібно наголосити, що результат оцінки за факторами залежить від того, як згруповані ці дослідження в статистичному комплексі.

Необхідно вказати й на обмеження у визначенні оцінки вірогідності впливу факторів. Якщо величина вирахованого коефіцієнта F_p перевищує його табличне значення F_T , то вплив досліджуваного фактора вважається вірогідним, а якщо не перевищує межу своїх випадкових коливань, то фактор не є суттєвим і не впливає на результат. Отже, не слід поспішати з висновком, оскільки причиною його невизначеності є недостатня кількість одиниць у вибірці для його переконливого підтвердження, а не різкий вплив факторів.

Інколи величина F_p може виявитися меншою від свого табличного значення через недостатньо різкий вплив фактора, що вивчається, або через недостатню чисельність вибірки. Причиною може бути й те, що помилка кожного з показників, узятих окремо, досить велика внаслідок завищеної неоднорідності досліджуваних даних. Вели-

чину F_p (занижену) зумовлюють і властивості самих факторів, такі як функціональні й близькі до них зв'язки між факторами, використання в аналізі однорічних даних та ін. Як наслідок – показники значно відрізняються від 0 або від 1, що збільшує їх можливі випадкові коливання. Це відображається на величині їх помилки, а від останньої залежить значення розрахованого коефіцієнта F_p .

Поспішний висновок щодо несуттєвості впливу фактора може тільки гальмувати подальші пошуки. Можливо, цим і пояснюється переконання окремих дослідників відносно статистичної оцінки вірогідності дослідження взагалі. Недоведеність істотності впливу фактора повинна не стримувати, а, навпаки, стимулювати подальші пошуки поліпшення експерименту як у відношенні техніки обробки, так і в підборі самого матеріалу. У такому випадку одержані позитивні результати стають ще більш неспростовними.

Щоб у дисперсійному аналізі мати об'єктивні результати, необхідно дотримуватися певних правил побудови (організації) дисперсійних комплексів. Якщо поділити групи на підгрупи (градації) таким чином, що в кожній із них рівні показників виявляться близькими за величиною, а між групами різко відрізняються, то дисперсійний аналіз може призвести до негативної відповіді на питання про істотність досліджуваних факторів. Це є наслідком того, що в загальній кількості показників у середній групі буде багато таких із них, які мало відрізняються один від одного, що може нівелювати відмінності між іншими. Різкі ж відмінності між середніми групами ніби зникнуть у великій кількості подібних один до одного середніх.

Висновки. Аргументуючи сказане, потрібно підкреслити, що дана обставина, обмежуючи можливість застосування дисперсійного аналізу в техніці, біології, медицині тощо, не така вже й небезпечна в галузі економіки. Тут оцінка в загальному і в цілому всіх відмінностей у характеристиках одиниць спостереження майже не має сенсу. В економічному аналізі вкрай важлива оцінка відмінностей між кожною групою.

Об'єктивність вибору варіантів побудови дисперсійних комплексів орієнтує на випробування їх інформаційних масивів на відповідність наближення емпіричного розподілу (а це інтервальний ряд) теоретичному, тобто підпорядкованості Закону нормального розподілу Гаусса – Лапласа. Ступінь такого наближення визначається коефіцієнтами асиметрії (A_S) та гостровершинності (ексцесом) розподілу (E_x). Значення коефіцієнта асиметрії до $\pm 0,25$, а показника ексцесу до $\pm 0,4$ будуть свідчити про наближення емпіричного розподілу до нормального. При такому характері розподілу є об'єктивні підстави щодо надійності інформаційної бази обраного варіанту побудови дисперсійного комплексу.

Із факту наявності в дисперсійному методі аналізу недоліків не випливає, що потрібно якось обмежити застосування цього методу в економіці. Мова йде не про обмеження, а про правильне використання його в аграрно-економічних дослідженнях, оскільки даний метод тільки у вказаному випадку високоефективний. Науково обґрунтований підхід до застосування цього досить ефективного прийому кількісного виміру факторів в економічних моделях надає йому суттєві переваги з-поміж інших методів багатofакторного статистичного аналізу ефективності виробництва.