

СТАТИСТИЧНІ АСПЕКТИ АНАЛІЗУ ДИНАМІКИ ІНФІКОВАНИХ НА COVID-19 В УКРАЇНІ

Перхун Лариса Петрівна,

кандидат педагогічних наук, доцент,
доцент кафедри економіко-математичних
дисциплін та інформаційних технологій;
завідувач сектору дистанційного навчання;

Товмаченко Ніна Миколаївна,

кандидат технічних наук, доцент,
доцент кафедри економіко-математичних
дисциплін та інформаційних технологій;
Національна академія статистики, обліку та аудиту

Коронавірусну інфекцію COVID-19 вперше в Україні було діагностовано 3 березня 2020 року в Чернівцях. З того часу накопичено досить великий масив щоденних даних у розрізі областей за показниками: загально інфікованих, інфікованих за добу, загально смертельних випадків, смертельних випадків за добу, загально одужало осіб, одужало за добу, зараз хворіють (на поточну дату) [1]. Різностороннє вивчення накопиченого масиву даних може дати відповіді на найрізноманітніші питання.

Наразі в Україні, як і у світі взагалі, відбуваються численні дискусії щодо доречності тих чи інших протиепідемічних заходів, або навіть мітинги протесту проти запровадження деяких із них. Місцева влада України завдяки процесам децентралізації у більшості випадків самостійно приймає рішення про способи запобігання поширенню інфекції та терміни їх застосування на підвідомчій території. Бажано мати інструмент, який дозволить працівникам органів місцевого самоврядування оцінювати ефективність запровадження конкретного заходу в схожих умовах.

Візуалізація ранжируваного ряду кількості інфікованих в Україні за добу станом на 26.10.2020 за областями України свідчить про суттєву диференціацію (рис. 1). А чи є ця диференціація статистично значущою? Яким чином ми можемо відстежувати зміну (статистично значущу) кількості інфікованих як реакцію на запровадження певного протиепідемічного заходу? Чи в усіх випадках перенесення позитивного досвіду певних територій на інші призведе до аналогічного результату?

Для отримання часткової відповіді на поставлені питання авторами обрано метод багатofакторного, зокрема двофакторного, дисперсійного аналізу з таких міркувань: по-перше, цей метод дозволяє дослідити вплив на числову результативну ознаку факторів, що вимірюються за номінальною або порядковою шкалою; по-друге, він дає можливість оцінити взаємодію факторів.

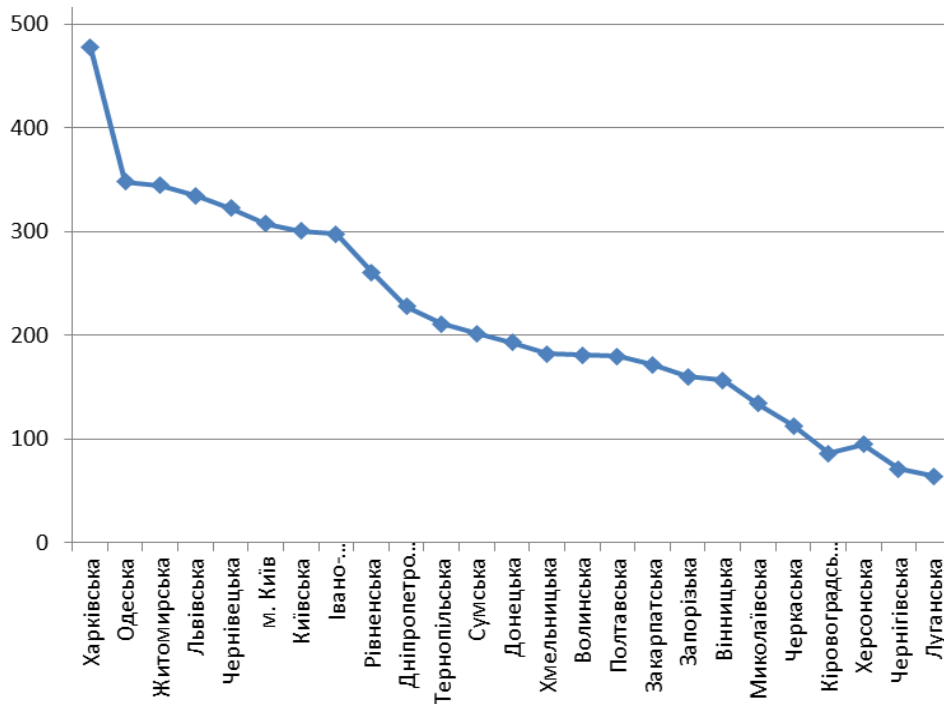


Рис. 1. Кількість інфікованих COVID-19 за добу за областями України станом на 26.10.2020 р.

За результативний показник (Y) було обрано кількість інфікованих за добу в розрізі областей. Перший фактор (фактор A) – області України. Кількість рівнів першого фактора $p=25$. Другим фактором, мали б бути конкретні протиепідемічні заходи, що впроваджені у кожній з областей. Але таких статистичних даних у загальному доступі немає. Проте в кожній області певних заходів дотримуються. Про їх ефективність опосередковано можна судити за кількістю інфікованих протягом певного періоду часу. Таким періодом визначено 3 тижні – інкубаційний період від 3 до 14 днів, плюс запізнення в обробленні тестів на COVID-19 від 5 до 7 днів. Отже, за другий фактор вирішено обрати номер тижня з трьох, взятих один за одним. Кількість рівнів другого фактора $q=3$. Кількість повторних вимірювань $r=7$ (за днями тижня).

Сутність методу двофакторного дисперсійного аналізу полягає в тому, щоб розкласти варіацію залежної змінної на частини, які відповідають окремому та спільному впливу на неї незалежних факторів.

Правило розкладання загальної дисперсії:

$$\sigma_{заг}^2 = \sigma_a^2 + \sigma_b^2 + \sigma_{ab}^2 + \sigma_\varepsilon^2, \quad (1)$$

де $\sigma_{заг}^2$ – загальна дисперсія;

σ_a^2 – дисперсія, обумовлена дією фактора A ;

σ_b^2 – дисперсія, обумовлена дією фактора B ;

σ_{ab}^2 – дисперсія обумовлена взаємодією факторів A та B ;

σ_{ε}^2 – дисперсія похибки вимірювання.

Потім, на основі статистичних критеріїв, зокрема критерію Фішера, перевіряється низка гіпотез:

- 1) відсутність ефекту фактора *A* (H_0 : диференціація кількості інфікованих у розрізі областей не є статистично значущою);
- 2) відсутність ефекту фактора *B* (H_0 : диференціація кількості інфікованих у розрізі трьох тижнів, що обрані послідовно, не є статистично значущою);
- 3) відсутність ефекту взаємодії факторів *A* та *B* (H_0 : взаємний вплив факторів *A* та *B* на кількість інфікованих за добу є статистично незначущим).

Розрахунки проводились на базі щоденної інформації про кількість інфікованих в Україні за добу в розрізі областей за три послідовних тижні – з 26.10.2020 по 15.11.2020 [1].

Результат розрахунків за методом двофакторного дисперсійного аналізу в пакеті SPSS (міжгрупові ефекти) представлено в табл. 1.

Таблиця 1

Фрагмент результату розрахунків за методом двофакторного дисперсійного аналізу в пакеті SPSS (тест міжгрупових ефектів)

Tests of Between-Subjects Effects
Dependent Variable: Y inf day

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	69771680,190	1	69771680,190	53,108	,000
	Error	7764318,270	5,910	1313780,839(a)		
Factor_A	Hypothesis	14541152,857	24	605881,369	16,805	,000
	Error	1730563,760	48	36053,412(b)		
Factor_B	Hypothesis	1487905,764	2	743952,882	20,635	,000
	Error	1730563,760	48	36053,412(b)		
Factor_A * Factor_B	Hypothesis	1730563,760	48	36053,412	4,138	,000
	Error	3920745,429	450	8712,768(c)		

У таблиці тесту міжгрупових ефектів (табл. 1) містяться значення сум квадратів і середніх квадратів для факторів *A*, *B* та їх взаємодії, розрахункові значення критерію Фішера. Для підтвердження чи спростування висунутих гіпотез аналізуються числа стовпчика «Sig». Гіпотеза H_0 з довірчою ймовірністю $P > 0,95$ відхиляється, якщо у відповідному рядку $\text{sig} \leq 0,05$.

За даними табл. 1 усі три нульові гіпотези відхиляються, тобто статистично значущими є окремі фактори *A* і *B*, а також їх взаємодія.

Отже, ми можемо стверджувати, що диференціація кількості інфікованих за добу в розрізі областей або послідовних тижнів є статистично значущою. Оцінювати результат запровадження певного протиепідемічного заходу можна через ефект взаємодії факторів *A* і *B*.

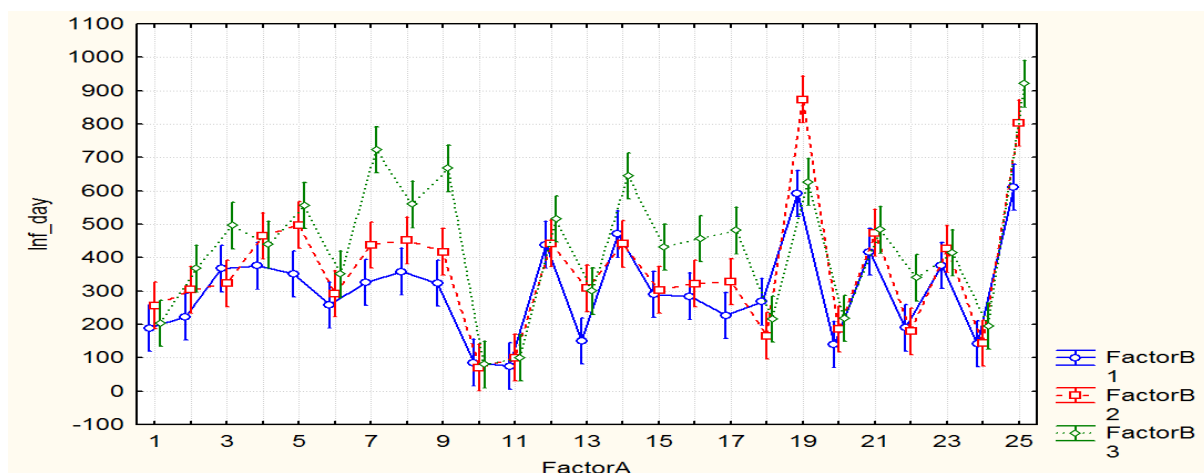
Результат розрахунку складових загальної дисперсії (1) та оцінювання впливу факторів *A*, *B* та їх взаємодії на залежну змінну *Y* (відсотку внеску складових у загальну дисперсію) показано в табл. 2.

Розрахунок складових загальної дисперсії

Джерело дисперсії	К-сть ступенів свободи	Сума квадратів SS	Середній квадрат MS	Розрахункове значення дисперсії	% дисперсії
Фактор А	$\gamma_1=24$	14541152,85	605881,36	$\sigma_a^2=33709,49$	67,98
Фактор В	$\gamma_2=2$	1487905,76	743952,88	$\sigma_b^2=3256,159754$	6,57
Взаємодія факторів АВ	$\gamma_3=48$	1730563,76	36053,41	$\sigma_{ab}^2=3905,80$	7,88
Похибка	$\gamma_4=450$	3920745,42	8712,76	$\sigma_\varepsilon^2=8712,76$	17,57
				$\sigma_{заг}^2=49584,23$	100%

Отже, результати двофакторного дисперсійного аналізу (табл. 1 і 2) статистично достовірно показують, що динаміка кількості заражених за добу в Україні пояснюється дією фактора *A* (область) на 67,8%, дією фактора *B* (номер тижня з трьох послідовно обраних) на 6,57% , взаємодією факторів *A* і *B* на 7,88%, іншими чинниками – на 17,57%.

Ефект взаємодії факторів *A* і *B* пропонується проаналізувати за середніми значеннями кількості інфікованих за тижнями і областями (дані візуалізовано на рис. 2).



Ряд 1 – тиждень 1 (суцільна лінія), ряд 2 – тиждень 2 (пунктирна лінія), ряд 3 – тиждень 3 (лінія крапками)

1 – Вінницька, 2 – Волинська, 3 – Дніпропетровська, 4 – Донецька, 5 – Житомирська, 6 – Закарпатська, 7 – Запорізька, 8 – Івано-Франківська, 9 – Київська, 10 – Кіровоградська, 11 – Луганська, 12 – Львівська, 13 – Миколаївська, 14 – Одеська, 15 – Полтавська, 16 – Рівненська, 17 – Сумська, 18 – Тернопільська, 19 – Харківська, 20 – Херсонська, 21 – Хмельницька, 22 – Черкаська, 23 – Чернівецька, 24 – Чернігівська, 25 – м. Київ.

Рис. 4. Взаємодія факторів *A* та *B*

За результатами графіків, поданих на рис. 2, майже у всіх областях простежується постійне зростання кількості інфікованих на COVID-19. Причому в деяких – досить швидко (Запорізька, Київська, Одеська, Харківська області, м. Київ), в інших – повільніше (Вінницька, Волинська, Дніпропетровська тощо).

Однак існують і такі області, де зміни відбулися в межах довірчих інтервалів середніх значень інфікованих за тиждень. Поширення такої позитивної практики на інші територіально-адміністративні одиниці дозволило б зменшити навантаження на систему охорони здоров'я України та зберегло б життя багатьом громадянам.

Подальший статистичний аналіз має виявити інші суттєві фактори, що впливають на кількість інфікованих в Україні: перелік протиепідемічних заходів, запроваджених в окремих областях, та культура їх дотримання; щільність населення; стан системи охорони здоров'я та ін.

Список використаних джерел

1. Коронавірус в Україні. Мінфін України. URL: <https://index.minfin.com.ua/reference/coronavirus/ukraine/>.
2. IBM SPSS Statistics. URL: <https://www.ibm.com/products/spss-statistics>.
3. StatSoft – Академия анализа данных. URL: <http://statsoft.ru/>.

ДОСЯГНЕННЯ ПРИБУТКОВОСТІ ТА ЕКОНОМІЧНОЇ БЕЗПЕКИ ПІДПРИЄМСТВА НА ОСНОВІ УПРАВЛІНСЬКОГО ОБЛІКУ

Пилипенко Олексій Іванович,

доктор економічних наук, доцент,
Національна академія статистики, обліку та аудиту

Сучасні умови господарювання диктують потребу в сукупності заходів, спрямованих на поліпшення управління господарською діяльністю. Залежність ефективності діяльності від управління нею зумовлює необхідність забезпечення прийняття обґрунтованих управлінських рішень з метою підвищення конкурентоспроможності підприємства. Ускладнення організаційних структур підприємств, посилення конкуренції, необхідність гнучкої системи управлінського обліку визначають затребуваність облікового забезпечення управління вітчизняними підприємствами. Актуальним постає питання забезпечення прибутковості та економічної безпеки підприємства на основі відповідної концепції планування, бухгалтерського обліку, контролю та економічного аналізу.

Обґрунтуванням концептуальних положень управлінського обліку займалися вчені А. Апчерч [1], М. Вахрушина [2], Н. Врублевский [3], К. Друри [4], Н. Кондраков [5], Б. Нидлз [6], М. Чумаченко [7], М. Щирба [8] та ін.

З розвитком суспільного виробництва бухгалтерський облік еволюціонував як інструмент управління господарською діяльністю і як джерело інформації для управління. Інформаційне забезпечення управління та якість облікової інформації тісно пов'язані з проблемою захисту інформації та системою економічної безпеки