

## СЕКЦІЯ 4. ВИКЛИКИ ІНФОРМАЦІЙНОГО ЗАБЕЗПЕЧЕННЯ СТАТИСТИКИ В УМОВАХ ВОЄННОГО СТАНУ

*Вдовин М. Л.,  
кандидат економічних наук,  
доцент, доцент кафедри статистики;  
Міщук Т. В.,  
здобувач другого (магістерського) рівня вищої освіти;  
Львівський національний університет імені Івана Франка*

### КОРЕЛЯЦІЙНО-РЕГРЕСІЙНИЙ АНАЛІЗ ВПЛИВУ ЗОВНІШНЬОЕКОНОМІЧНИХ ЧИННИКІВ НА ДОБРОБУТ НАСЕЛЕННЯ УКРАЇНИ

Добробут населення є досить абстрактною величиною, яка формується під впливом значної кількості чинників. Проте найбільш повно, на нашу думку, його характеризує показник ВВП на одну особу, як показник економічного зростання країни та ідентифікатор виміру рівня життя населення.

Проблемам дослідження добробуту населення та сталого розвитку регіонів України присвячена низка наукових праць. Зокрема, у [1] описано просторово-панельні методи моделювання сталого розвитку, у [2,3] - економетричні методи аналізу макроекономічних процесів та дослідження зовнішньоекономічної діяльності, а у [4] статистичні методи аналізу добробуту населення України. Велику увагу також зосереджено на показниках зовнішньої торгівлі як однієї зі складових частин добробуту. Так, у [5] розглянуто ЗТД України з країнами ЄС, а у [6] ігрове моделювання зовнішньоекономічної діяльності країн Центральної та Східної Європи.

Для здійснення моделювання розглянемо як результативну ознаку ( $Y_{\text{fact}}$ ) саме ВВП на 1 особу, а факторними будуть - середньорічна заробітна плата ( $X_1$ ), витрати на придбання товарів та послуг ( $X_2$ ), суспільні витрати на освіту ( $X_3$ ), суспільні витрати на охорону здоров'я ( $X_4$ ), чисельність населення із середньодушовими доходами у місяць, нижчими законодавчо встановленого прожиткового мінімуму ( $X_5$ ), середня очікувана тривалість життя в Україні ( $X_6$ ), прямі іноземні інвестиції ( $X_7$ ), імпорт товарів та послуг ( $X_8$ ). Серед зовнішньоекономічних чинників обрано прямі іноземні інвестиції та імпорт товарів та послуг. Дані взято з офіційної сторінки Державної служби статистики України і систематизовано у табл. 1.

Таблиця 1

## Показники добробуту населення України у 2006-2021 рр.

Період, роки	$Y_{\text{факт}}$ (грн)	$X_1$ (млрд грн)	$X_2$ (млрд грн)	$X_3$ (млрд грн)	$X_4$ (млрд грн)	$X_5$ (%)	$X_6$ (років)	$X_7$ (млрд дол США)	$X_8$ (млрд дол. США)
2006	12076	205.1	385.7	49.78	32.78	21.4	68.1	21.61	48.8
2007	16150	278.9	509.5	53.47	42.85	12.7	68.25	29.54	65.6
2008	21419	366.4	695.6	70.38	59.62	7.1	68.27	35.62	92
2009	20564	376.1	712.7	77.41	68.38	5.8	69.29	40.05	80.6
2010	24429	459.2	838.2	91.1	80.83	8.6	70.44	44.81	81.1
2011	29980	521.07	1024.25	97.60	91.27	7.8	71.02	50.33	88.86
2012	32480	609.39	1194.79	111.18	104.30	9	71.15	55.30	91.36
2013	33965	630.73	1304.03	115.96	111.14	8.3	71.37	57.06	83.35
2014	36904	615.02	1316.76	109.52	113.24	8.6	71.37	45.92	84.75
2015	46413	709.59	1568.17	127.12	149.16	6.4	71.38	45.42	86.98
2016	55899	898.33	1840.26	139.97	174.48	3.8	71.68	47.16	80.57
2017	70170	1209.10	2359.99	182.98	214.35	2.4	71.98	47.02	75.10
2018	84228	1529.37	2884.97	214.00	266.16	1.3	71.76	46.30	63.49
2019	94633	1758.59	3356.99	241.41	282.14	1.1	72.01	52.46	67.74
2020	101138	1832.29	3520.55	252.88	318.88	1.6	71.35	51.18	71.32
2021	131944	2133.65	4324.19	267.7	343.67	1.3	71.42	62.42	72.82

Джерело: побудовано автором за даними [7].

Лінійне рівняння залежності ВВП на одну особу України від чинників  $X_1$ - $X_8$  має вигляд:

$$Y_{\text{теор}} = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 + a_5X_5 + a_6X_6 + a_7X_7 + a_8X_8 + e \quad (1)$$

де  $a_i$ ,  $i=0,1,\dots,8$  – коефіцієнти лінійної регресії;  $Y_{\text{теор}}$  – розрахунковий обсяг ВВП на одну особу країни;  $X_1$  – середньорічна заробітна плата,  $X_2$  – витрати на придбання товарів та послуг,  $X_3$  – суспільні витрати на освіту,  $X_4$  – суспільні витрати на охорону здоров'я,  $X_5$  – чисельність населення із середньодушовими доходами у місяць, нижчими законодавчо встановленого прожиткового мінімуму,  $X_6$  – середня очікувана тривалість життя в Україні,  $X_7$  – прямі іноземні інвестиції,  $X_8$  – імпорт товарів та послуг,  $e$  – випадкова величина.

Досить частою є проблема мультиколінеарності регресійної моделі, тому з метою виявлення прихованих залежностей між факторними ознаками проведемо їх кореляційний аналіз (табл.2).

Таблиця 2

**Матриця парних кореляцій факторних ознак добробуту населення**

	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$
$X_1$	1	0.99755	0.99451	0.99376	-0.77634	0.65861	0.63761	-0.04841
$X_2$	0.99755	1	0.99239	0.99399	-0.77471	0.67400	0.65761	-0.05703
$X_3$	0.99451	0.99239	1	0.99591	-0.80149	0.71977	0.67429	-0.03763
$X_4$	0.99376	0.99399	0.99591	1	-0.79954	0.70261	0.64454	-0.09492
$X_5$	-0.77634	-0.77471	-0.80149	-0.79954	1	-0.71839	-0.71285	-0.06357
$X_6$	0.65861	0.67400	0.71977	0.70261	-0.71839	1	0.82862	0.08237
$X_7$	0.63761	0.65761	0.67429	0.64454	-0.71285	0.828616	1	0.42649
$X_8$	-0.04841	-0.05703	-0.03763	-0.09492	-0.06357	0.08237	0.42649	1

Джерело: розроблено автором за даними [7].

Як бачимо з даних, які репрезентовані у матриці, не всі обрані для кореляційно-регресійного аналізу фактори між собою мають високу щільність. Це стосується факторів  $X_5$ ,  $X_6$ ,  $X_7$ ,  $X_8$ . Натомість, коефіцієнт кореляції між факторами  $X_1$ ,  $X_2$ ,  $X_3$  та  $X_4$  наближений до одиниці, а це означає, що між ними існує тісний взаємозв'язок.

Також було проведено перевірку на наявність мультиколінеарності за методом Фаррара-Глобера. Результати перевірки факторних ознак на мультиколінеарність за  $\chi^2$ -«хі-квадрат» показали, що в масиві незалежних змінних мультиколінеарність присутня, так як виконується основна нерівність. Хі-квадрат фактично становить 246.247, а табл 41,3. За результатами перевірки факторних ознак на мультиколінеарність за F-критерієм, усі незалежні змінні, крім останньої, є мультиколінеарними з іншими. Результати перевірки факторних ознак на мультиколінеарність за t-критерієм Стюдента показали з отриманих результатів, між деякими парами факторних ознак ( $X_5$  і  $X_2$ ,  $X_6$  і  $X_3$ ,  $X_7$  і  $X_3$ ) існує мультиколінеарність. Для позбавлення від даного ефекту необхідно виключити одну із мультиколінеарних змінних, у нашому випадку це  $X_3$  – суспільні витрати на освіту, і провести процедуру знову.

За використанням цієї методики було встановлено, що на обсяг ВВП на одну особу України суттєво впливає лише один показник – імпорт товарів та послуг ( $X_8$ ). Перевірка моделі на адекватність дає можливість можна стверджувати, що модель є адекватною, оскільки значимість F-критерію та P-значення менше ніж 0,05. Отже, показник  $X_8$  є статистично значимим. За результатами проведеного дослідження, лише він, а також показники  $X_3$  та  $X_4$  мали P-значення менше 0,05. Проте, похибки останніх були більше 1. Тоді як для  $X_1$  P-значення було 0,48,  $X_2$  – 0,82, для  $X_5$  – 0,52, для  $X_6$  – 0,68, для  $X_7$  – 0,95. Виходячи з наведеного, зі зовнішньоекономічних чинників лише вплив імпорту товарів та послуг на ВВП на одну особу України є статистично обґрунтованим.

Рівняння досліджуваної залежності буде мати вигляд:

$$Y_{teor} = 245,65X_8, \quad (2)$$

Адекватність моделі було визначена за допомогою розрахунку множинного коефіцієнта детермінації, який оцінює частку варіації

результату за рахунок представлених у рівнянні факторів у загальній варіації результату. У даному випадку, коефіцієнт детермінації  $R^2 = 0,914$ . Він є близьким до одиниці. Це свідчить про щільний зв'язок між розрахунковою величиною ВВП на одну особу та обсягом імпорту товарів та послуг. Також, значення  $R^2$  говорить про те, що 91,4 % зміни обсягу імпорту товарів та послуг ( $X_8$ ) описували зміни обсягу ВВП на душу населення ( $Y_{\text{fact}}$ ). Стандартна похибка показника  $X_8$  становить 0,67, тому коефіцієнт  $X_8$  є значимим.

Також варто звернути увагу на значення F-критерію Фішера та t-критерій Стьюдента, які визначають значимість та адекватність моделі. Оскільки емпіричне значення критерію Фішера більше від його критичного значення  $F_{\text{ем}}=3154,76 > F_{\text{кр}}=4,6$ , то нульову гіпотезу відхиляємо і з довірчою ймовірністю  $p=0,95$  вважаємо, що побудована модель адекватна реальній дійсності. Аналогічно емпіричне значення t-критерій Стьюдента, що дорівнює 47,86 більше за критичне – 2,13, то нульову гіпотезу відхиляємо і з довірчою ймовірністю  $p=0,95$  вважаємо, що кореляційна залежність величини імпорту товарів та послуг від величини ВВП на одну особу є статистично значущою.

Застосування методологічних основ статистичного моделювання чинників економічного зростання та рівня добробуту дозволяє сформулювати висновок, що метод кореляційно-регресійного аналізу дає можливість отримувати достовірні репрезентативні висновки при наявності навіть нечисленних сукупностей. Тому, з метою підвищення рівня добробуту населення слід звернути увагу на збереження позитивної тенденції до покращення досліджуваних показників у наступних роках, впливаючи на них за допомогою відповідних важелів державної політики.

### **Список використаних джерел**

1. Зомчак Л., Коваль Л. Сталий розвиток регіонів України: просторово-панельний підхід. Інфраструктура ринку. 2022. Вип. 65. С. 211–215
2. Зомчак Л.М., Рудницька Я.І. Економетричне моделювання залежностей між макроекономічним станом України та обсягами залучених депозитів. Причорноморські економічні студії. 2020. Вип. 49. С. 196–200
3. Пласконь С., Сенів Г., Новосад І., Маслій В. Застосування економетричного моделювання в оцінюванні зовнішньоекономічної діяльності України. Економічний аналіз. 2020. Том 30. № 3. С. 25-32. DOI: <https://doi.org/10.35774/econa2020.03.025>
4. Вдовин М.Л., Зомчак Л.М., Міщук Т.В. Аналітика сучасного стану добробуту населення України. Економіка та суспільство. 2022. Вип. 39. <https://doi.org/10.32782/2524-0072/2022-39-77>
5. Вдовин М. Л., Хромова М. В. Проблеми та перспективи розвитку зовнішньоторговельних відносин України з країнами ЄС. Глобальні та національні проблеми економіки. 2015. Вип. 6. С. 61–63.

6. Vdovyn, M., Zomchak, L., Panchyshyn, T. Modeling of economic systems using game theory. *Věda a perspektivy*. 2022. 7 (14), 91-100. DOI: [https://doi.org/10.52058/2695-1592-2022-7\(14\)-91-100](https://doi.org/10.52058/2695-1592-2022-7(14)-91-100). URL: <http://perspectives.pp.ua/index.php/vp/article/view/2034/2034>

7. Державна служба статистики України. Офіційний сайт. URL: <https://www.ukrstat.gov.ua/>(дата звернення: 07.03.2023).

*Горобець Олена Олександрівна,  
кандидат економічних наук, доцент кафедри статистики,  
інформаційних технологій та математичних методів в економіці;  
Національна академія статистики, обліку та аудиту*

## **ПЕРСПЕКТИВИ ІМПЛЕМЕНТАЦІЇ СЕНТИМЕНТ-АНАЛІЗУ В ОФІЦІЙНУ СТАТИСТИКУ**

Війна в Україні змушує статистиків вдаватися до пошуку та імплементації нових, альтернативних методів збирання та аналізу даних. Враховуючи потужний інформаційно-технологічний розвиток, з'являється можливість використовувати й технології штучного інтелекту.

Оброблення природньої мови, як галузі штучного інтелекту, має у своєму арсеналі потужний інструментарій методів, які допомагають розуміти, інтерпретувати та генерувати людську мову в зрозумілому, як комп'ютерам так і людям, способі. Одним із методів оброблення природньої мови є сентимент-аналіз, перевагою якого є визначення емоційного забарвленості тексту. За допомогою цього методу є можливість автоматизовано аналізувати великі обсяги текстової інформації та зрозуміти ставлення суспільства до певних подій, продуктів та ін.

Збирання даних для сентимент-аналізу може відбуватися з різних джерел в залежності від мети аналізу, зокрема із соціальних медіа, коментарів або відгуків на веб-сайтах, за допомогою опитувань та форм зворотного зв'язку, аналізу відео-, аудіо- та графічного вмісту, а також новин, статей.

Сентимент-аналіз має своєрідну систему методів, використання яких залежить від завдання, обсягу даних та вимог до аналізу.

Загалом методи, які формують своєрідну систему можна розділити на три категорії:

Перша категорія: застосовуються певні правила в рамках визначення настрою, ґрунтуючись на емоційне забарвлення окремих слів у тексті. Зокрема у цій категорії методів використовують лексикографічні джерела.

Друга категорія: використовується байєсівський підхід для прогнозування ймовірності приналежності тексту до «позитивного», «негативного» або ж «нейтрального» класів; задля визначення меж між зазначеними класами, зазвичай, використовують метод опорних векторів; з метою визначення взаємозв'язків у текстах найчастіше використовують нейронні мережі.