

ДИНАМІЧНА КОРЕЛЯЦІЙНА МОДЕЛЬ ЕЛЕКТОРАТНИХ УПОДОБАНЬ ВИБОРЦІВ НА ПРИКЛАДІ ВИБОРІВ 2010 РОКУ

Селезньова Н.П.

кандидат фізико-математичних наук, доцент кафедри математичної фізики;

Бондаренко М.А.

магістрантка ФСП 6 курсу НТУУ «КПІ»

Авторами проведено, напередодні виборів у Президенти України 2010 року, польове дослідження вподобань виборців шляхом анкетного опитування 150 студентів НТУУ „КПІ” протягом трьох періодів (трьох хвиль). Результати опитування досліджено методами кореляційного аналізу, при цьому обчислено та систематизовано такі показники, як: вибіркові коефіцієнти кореляції, коефіцієнти кореляції Спірмена, множинний і частковий коефіцієнти кореляції та детермінації. Зроблено висновки з отриманих показників.

The authors conducted a field study of voters preferences by questionnaire of 150 students KPI for three periods (three waves) on the eve of a Presidential election in 2010. The results of the questioning were studied by correlation analysis, while evaluated and systematized such factors as: sample correlation coefficient, Spearman's correlation coefficients, multiple and partial correlation coefficients and determination. The conclusions from these figures.

Авторами проведено, накануне выборов в Президенты Украины 2010 года, полевое исследование предпочтений избирателей путем анкетного опроса 150 студентов НТУУ „КПИ” на протяжении трех периодов (трех волн). Результаты опроса исследованы методами корреляционного анализа, при этом вычислено и систематизировано такие показатели, как: выборочные коэффициенты корреляции, коэффициенты корреляции Спирмена, множественный и частичный коэффициенты корреляции и детерминации. Сделаны выводы из полученных показателей.

Ключові слова: вибори, кореляція.

Колись знаменитий Гаус сказав: «Математика – це цариця наук»; у наш час стало очевидним те, що математика має більш високий статус: вона є служницею всіх наук, допомагаючи їм та надаючи їм адекватний апарат для опису різноманітних фактів і явищ. Більш того, математика - це та служниця, без якої і пани не є панами, тобто без якої науку і за науку визнати неможливо, бо «рівень науковості» тої чи іншої дисципліни можна виміряти об'ємом застосованих у ній математичних міркувань, глибиною і змістом характерних для цієї дисципліни дедуктивних висновків.

Вважається що демократія – це мистецтво політичного маневрування. Ця теза є вірною далеко не у всьому. Мета запропонованої роботи – показати, що і математика грає далеко не останню роль у політиці.

У нашій роботі такий розділ математики як статистика надає нам «універсальний ключ», який є придатним для відмикання усіх замків, навіть таких складних що виникають у людських стосунках, а саме у політиці.

Статистичний аналіз має багатовікову історію, по-справжньому оформлення прикладної статистики як методології роботи з числовими даними можна пов'язувати з ім'ям Ф. Гальтона, який в кінці XIX ст. вперше застосував статистичний аналіз в біології та психології. У 1888 р.

вчений виступив з доповіддю на засіданні Королівського товариства «Кореляції і їх вимір, переважно за антропометричними даними».

Сьогодні західні дослідники активно застосовують статистичні методи різного ступеня складності, про які наші учені дізнаються або з публікованих статей, або з виступів на конференціях. За останні роки було видано достатньо багато монографій і навчальних посібників, написаних російськими та вітчизняними авторами. Але вони охоплюють в основному прості і найпоширеніші способи статистичного аналізу: описову статистику, зіставлення двох вибірок, а все, що виходить за рамки цього обов'язкового мінімуму розглядається тільки в ознайомлювальному форматі (А.Д. Наслідов; А. П. Кулаїчев) або присвячено окремим методам (А.Н.Гусев, О.В.Мітіна, І. Б. Михайлівська). В результаті велика кількість методів, вже широко відомих і активно використовуваних колегами за кордоном, що дозволяють перевіряти цікаві і диференційовані гіпотези, до теперішнього часу відомі вузькому колу дослідників і застосовуються вкрай рідко (наприклад, дискримінантний, кореляційний аналіз, аналіз коваріацій) або не використовуються взагалі (логлінійний аналіз).

У нашій роботі для дослідження електоральних настроїв ми застосуємо кореляційний аналіз. Саме за допомогою

кореляційного аналізу побудуємо таку систему показників, яка допоможе розібратись політичним аналітикам у тонкощах передвиборчої ситуації.

Величина кореляції визначається коефіцієнтом кореляції; цей показник широко застосовується і в наші дні, його можна знайти у будь-якому підручнику зі статистики, соціології чи біометрії. Окрім коефіцієнта, який ввів до розгляду Гальтон зараз широко застосовуються і інші коефіцієнти кореляції для різних класів задач.

Кореляція („co-relation“) означає статистичну „співзалежність“ змінних. Систематичне вивчення показників кореляції було започатковано в роботах учня Гальтона Карла Пірсона приблизно в кінці XIX століття [3, с. 30]. Сьогодні дану методологію використовують...

В нашій роботі ми вводимо до розгляду систему показників кореляції, які допомагають вивчити та систематизувати зміну настроїв електорату під час передвиборчої кампанії.

На сьогоднішній день одним із головних факторів, які здійснюють вплив на державну політику, є електоральна поведінка.

Електоральною поведінкою називається сукупність дій громадян,

пов'язаних із здійсненням місцевих або загальнонаціональних виборів в органи влади, а також їх участю в референдумах.

Саме вибори представницьких органів влади актуалізують потребу в з'ясуванні електоральних настроїв, адже це надасть можливість для розробки стратегії проведення виборчої кампанії та здійснення управління нею.

Для збору первинних даних авторами застосовувався метод анкетного опитування, яке було проведено в три етапи з 14.12.09 по 18.12.09 (1 хвиля), з 21.12.09 по 26.12.09 (2 хвиля) та з 11.01.10 по 16.01.10 (3 хвиля), що в подальшому дало можливість прослідкувати зміни настроїв електорату.

Перед респондентами ставилося завдання: надати рейтингові бали («0» - відсутність інформації про кандидата, «1» - повне «несприйняття» кандидата, наступні бали в порядку зростання надаються «кращим» кандидатам, 9-абсолютна підтримка респондентами) дев'ятьом кандидатам у Президенти та дев'яти відповідним політичним партіям (блокам). Генеральну сукупність складають студенти від 18 років різних факультетів НТУУ «КПІ» денної форми навчання. Вибіркова сукупність складається з 150 студентів.

У результаті опитування були отримані наступні дані:

Таблиця 1

Кількість балів, які отримали кандидати у Президенти

	Гриценко	Литвин	Симоненко	Тіріпко	Тимошенко	Тягнибок	Ющенко	Янукович	Яценюк
1 хвиля	339	451	334	515	539	329	363	352	454
2 хвиля	316	367	340	645	563	275	382	419	465
3 хвиля	366	469	339	584	663	308	371	410	536

Таблиця 2

Кількість балів, які отримали партії(блоки)

	Блок Анатолія Гриценка	Блок Литвина	КПУ	"Сильна Україна"	БЮТ	"Свобода"	Наша Україна	Партія регіонів	"Фронт змін"
1 хвиля	335	419	353	516	533	334	400	355	421
2 хвиля	314	355	336	457	488	320	390	426	424
3 хвиля	364	438	342	505	613	337	386	413	340

Подальшу обробку отриманої інформації здійснено за допомогою методів математичної статистики, а саме кореляційного аналізу. Даний метод вказує наявність залежності між досліджуваними ознаками, її величину та напрям.

При вивченні уподобань виборців їм було запропоновано надати певні бали кандидатам та партіям (блока), тобто ми маємо справу з ординальними

(порядковими) змінними, які вимірюються у порядковій шкалі. У цій шкалі можна встановити лише порядок, в якому об'єкти вибудовуються за ступенем прояви ознаки. Адже якщо одного кандидата виборець оцінив, наприклад, у п'ять балів, а іншого у вісім, то не можна стверджувати наскільки або у скільки разів один кандидат кращий за іншого. Лише можна стверджувати, що один із кандидатів є кращим за іншого. У такому

випадку проблема оцінки тісноти зв'язку вирішується шляхом впорядкування або ранжування об'єктів аналізу за ступенем означеності вимірюваних ознак. При цьому кожному об'єкту присвоюється певний номер, який називають рангом. У нашому випадку присвоїмо ранг „1” об'єкту з найбільшим значенням ознаки, наступному за ним – ранг „2” і т.п.

За міру зв'язку візьмемо коефіцієнт кореляції Спірмена, який визначається наступною формулою[1, с. 201]:

$$K_c = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n}, \quad (1)$$

де $d_i = x_i - y_i$, x_i, y_i -ранги, присвоєні ознакам, n - об'єм вибірки. .

При повному прямому зв'язку $K_c = -1$, в усіх інших випадках $|K_c| < 1$.

Застосування рангового коефіцієнту кореляції Спірмена обумовлено тим, що він є непараметричним аналогом параметричних методів кореляції і є зручним для визначення ступені зв'язку між двома змінними, значення яких представлено рангами, а не стандартизованими оцінками. Логічне обґрунтування застосування коефіцієнта Спірмена не вимагає дотримання строго визначеного набору припущень, а, отже, він є непараметричною статистикою.

Обчислимо коефіцієнти кореляції Спірмена за формулою (1) між першою та другою $K_{C(12)}$, першою та третьою $K_{C(13)}$, другою та третьою $K_{C(23)}$ хвилями. Для обчислень побудуємо наступні таблиці:

Таблиця 3

	1 хвиля	Ранги	2 хвиля	Ранги
Тимошенко	539	1	316	8
Тігіпко	515	2	465	3
Яценюк	454	3	340	7
Литвин	451	4	645	1
Ющенко	363	5	367	6
Янукович	352	6	275	9
Гриценко	339	7	419	4
Симоненко	334	8	563	2
Тягнибок	329	9	382	5

$$K_{C(12)} = 1 - \frac{6 \cdot 12}{(9^3 - 9)} = 0,90$$

Таблиця 4

	1 хвиля	Ранги	3 хвиля	Ранги
Тимошенко	539	1	366	7
Тігіпко	515	2	536	3
Яценюк	454	3	339	8
Литвин	451	4	584	2
Ющенко	363	5	469	4
Янукович	352	6	308	9
Гриценко	339	7	410	5
Симоненко	334	8	663	1
Тягнибок	329	9	371	6

$$K_{C(13)} = 1 - \frac{6 \cdot 10}{(9^3 - 9)} = 0,92$$

Таблиця 5

	2 хвиля	Ранги	3 хвиля	Ранги
Литвин	645	1	584	2
Симоненко	563	2	663	1
Тігіпко	465	3	536	3
Гриценко	419	4	410	5
Тягнибок	382	5	371	6
Ющенко	367	6	469	4
Яценюк	340	7	339	8
Тимошенко	316	8	366	7
Янукович	275	9	308	9

$$K_{C(23)} = 1 - \frac{6 \cdot 10}{(9^3 - 9)} = 0,92$$

Як бачимо коефіцієнти кореляції Спірмена є досить високими, що свідчить про суттєвий зв'язок між вподобаннями виборців протягом періоду дослідження. Цей результат є досить стабільним, отже, виборці протягом досліджуваного періоду суттєво не змінювали своїх уподобань.

Дані про статистичну залежність зручно задавати також у вигляді кореляційної таблиці. Розглянемо залежність між вподобаннями виборців протягом першої, другої та третьої хвили за допомогою вибіркового коефіцієнту кореляції, який визначається такою формулою:

$$r_{ij} = \frac{\sum_{i=1}^l \sum_{j=1}^m (x_i - \bar{x})(y_j - \bar{y})n_{ij}}{ns_x s_y}, \quad (2)$$

де n - об'єм вибірки $\bar{x}, \bar{y}$ - середні значення ознак, x_i, y_j - варіанти відповідних вибірок, n_{ij} - відповідні частоти, s_x, s_y - відповідні середньоквадратичні відхилення.

Таблиця 6

Кореляційна матриця 1

	1	2	3
1	1	0,870642	0,973805
2		1	0,894892
3			1

Співставивши кореляційні матриці 1 та 2, бачимо, що зв'язок між хвилями по кандидатам у Президенти є вищим ніж зв'язок між хвилями по блокам чи партіям, а

Вибірковий коефіцієнт кореляції набуває значень на відрізку $[-1; 1]$, тобто $-1 \leq r_{ij} \leq 1$.

В залежності від того, наскільки $|r_{ij}|$ наближається до 1, розрізняють зв'язок слабкий, помірний, помітний, досить тісний, тісний та суттєво тісний, тобто чим ближче $|r_{ij}|$ до 1, тим міцнішим є зв'язок.

При $r_{ij} = \pm 1$ кореляційний зв'язок представляє лінійну функціональну залежність. При $r_{ij} = +1$ пряма лінійна залежність, при $r_{ij} = -1$ - обернена лінійна залежність.

За формулою (2) обчислимо елементи кореляційної матриці між трьома хвилями між кандидатами у Президенти за даними таблиці 1 (кореляційна матриця 1), між партіями (блоками) за даними таблиці 2 (кореляційна матриця 2).

Таблиця 7

Кореляційна матриця 2

	1	2	3
1	1	0,840826	0,865433
2		1	0,762302
3			1

це свідчить про більш слабкий зв'язок насторів виборців відносно певних кандидатур у Президенти та відповідних їм партій, що демонструє той факт, що

виборець частіше краще орієнтується на певну особистість ніж на партію, лідером якої є ця особистість.

Далі проведемо багатовимірний кореляційний аналіз, в якому ставляться та вирішуються дві типові задачі:

1) визначення тісноти зв'язку між однією із змінних та сукупністю інших змінних, які включені в аналіз;

2) визначення тісноти зв'язку між змінними при фіксуванні або виключенні впливу інших змінних.

Вище перелічені задачі вирішуються шляхом обчислення та аналізу множинного і часткового коефіцієнтів кореляції.

Множинний коефіцієнт кореляції (МКК) (для партій та між трьома хвилями).

Ми маємо три змінні – перша, друга та третя хвили ($p = 3$) (таб.2, 7). Отже, для обчислення множинного коефіцієнту кореляції застосуємо наступну формулу:

$$R_{i,jk} = \sqrt{\frac{r_{ij}^2 + r_{ik}^2 - 2r_{ij} \cdot r_{ik} \cdot r_{jk}}{1 - r_{jk}^2}}, \quad (3)$$

де $r_{ij} (i, j = 1, 2, \dots, p)$ -

МКК має межі $0 \leq R \leq 1$. Чим ближчим є R до 1, тим тіснішим є зв'язок.

За формулою (3), маємо:

Таблиця 8

$R_{1,23}$	0,91
$R_{3,12}$	0,87
$R_{2,13}$	0,71

Отримані результати свідчать про те, що є суттєвий зв'язок між вподобаннями виборців у першій хвилі та у другій і третій хвилі. Множинний коефіцієнт детермінації

$R_{1,23}^2 = 0.83$ показує, що варіація вподобань виборців у першій хвилі на 83% визначає вподобання виборців у другій та третій хвилях [2, с. 441]:.

	Тимошенко	Ющенко	Яценюк	Литвин	Симоненко	Янукович	Тягнибок	Гриценко	Тігіпко
БЮТ	0,69	0,50	0,26	0,36	0,22	0,03	0,27	0,28	0,04
Наша Україна	0,46	0,79	0,33	0,33	0,27	-0,02	0,37	0,32	0,07
Фронт змін	0,24	0,41	0,67	0,46	0,35	0,16	0,35	0,22	0,08
Блок Литвина	0,31	0,31	0,50	0,78	0,42	0,25	0,37	0,25	-0,08
КПУ	0,20	0,30	0,17	0,40	0,68	0,22	0,28	0,35	-0,04
Партія регіонів	0,11	0,05	0,12	0,22	0,16	0,82	0,12	0,12	0,07
Свобода	0,25	0,45	0,31	0,47	0,34	0,20	0,83	0,45	0,08

Частковий коефіцієнт кореляції (ЧКК) (для партій між трьома хвилями, за таб.2, 7)

Ми маємо три змінні – перша, друга та третя хвили ($p = 3$), отже, для обчислення часткового коефіцієнту кореляції використовуємо наступну формулу:

$$r_{ij.k} = \frac{r_{ij} - r_{ik} \cdot r_{jk}}{\sqrt{(1 - r_{ik}^2)(1 - r_{jk}^2)}}, \quad (4)$$

де $r_{ij} (i, j = 1, 2, \dots, p)$ - обчислюється за формулою (2). ЧКК може набувати значень на проміжку $[-1; 1]$.

Результати обчислення часткового коефіцієнту кореляції між трьома хвилями (для партій) за формулою (4) запишемо у вигляді таблиці 9.

Таблиця 9

$r_{12,3}$	0,15
$r_{31,2}$	0,64
$r_{23,1}$	0,13

Порівнюючи часткові коефіцієнти кореляції $r_{ij.k}$ з відповідними парними

коефіцієнтами r_{ij} (кореляційна матриця 2), бачимо, що завдяки «очищенню зв'язку»

найбільше змінився коефіцієнт r_{12} , отже між першою та другою хвилею є суттєвий кореляційний зв'язок, а третя хвиля має менш суттєвий вплив на результати вподобань виборців у другій та першій хвилях.

Дослідимо відповідність між вподобаннями виборців до кандидатів у Президенти та відповідними партіями шляхом обчислення коефіцієнтів кореляції за формулою (2). Занесемо отримані дані до таблиці 10 [1, с. 442]:.

Таблиця 10

Кореляційна матриця між вподобаннями виборців до кандидатів у Президенти та відповідними партіями

Блок									
Гриценка	0,17	0,27	0,21	0,17	0,34	0,08	0,44	0,81	0,13
Сильна Україна	-0,02	0,13	0,11	-0,02	0,09	0,09	0,07	0,14	0,98

З цієї таблиці випливає, що найбільш „свідомими” є виборці, які голосуватимуть за Тігіпка та партію „Сильна Україна”, бо коефіцієнт кореляції між ними дорівнює 0,98, а найменш „свідомими” – виборці кандидата Яценюка та відповідної партії „Фронт змін” , бо коефіцієнт кореляції між ними дорівнює 0,67. Найбільше „не люблять” одне одного виборці кандидата Тігіпка та тих, хто голосуватиме за „Блок Литвина” (коефіцієнт кореляції -0,08) і „КПУ” (коефіцієнт кореляції -0,04). Також антагоністами у своїх електорат них уподобаннях є виборці, що голосуватимуть за кандидата Тимошенко та партію „Сильна Україна” (коефіцієнт кореляції -0,02).

Загалом в даній роботі створено модель на основі систематизації різних показників кореляції. Моделі бувають змістовні та формальні. Ми запропонували змістовну модель. Загалом соціальні системи не просто функціонують у часі – вони ще й приймають рішення, здійснюють вибір шляху подальшого розвитку. Тому в даній роботі запропоновано системний підхід до вивчення електоральної поведінки шляхом збору та статистичної обробки первинної інформації отриманої анкетуванням певної групи виборців. Отже, авторами отримано такі результати:

Згідно таблиці 1 лідерами у передвиборчій боротьби за президентську посаду є Тимошенко, Тігіпка, Яценюк, а лідерами серед партій (блоків) – „БЮТ”, „Сильна Україна”, „Блок Литвина”.

За таблицею 6 ми бачимо, що найменше змінюють свою думку протягом дослідження (1,2,3 хвили) виборці за кандидатів у Президенти, а за партії (блоки) протягом трьох хвиль вподобання виборців більше змінювалися ніж за кандидатів у Президенти.

Як показує множинний коефіцієнт детермінації $R_{1.23}^2 = 0.83$, виборці визначилися з голосуванням у першій хвилі на 83% по партіям у порівнянні з другою та третьою хвилями.

З таблиці 10 бачимо повну невідповідність у вподобаннях виборців між партіями та їх лідерами, тобто „симпатія” виборців до лідера не завжди означає їх „симпатію” до відповідної партії.

Аналізуючи таблицю 10, бачимо, що виборці в основному визначилися з голосуванням протягом першої та другої хвилі.

Отже, застосована нами статистична методологія допомагає з’ясувати

електоратні уподобання, що представляють неабиякий інтерес для суб’єктів виборчого процесу - партій і кандидатів, адже без надійної аналітичної та інформаційної бази про передвиборчу ситуацію, обґрунтувати стратегію і тактику виборчої кампанії неможливо. Саме агрегація та артикуляція інтересів окремих страт, слоїв, класів дозволить кандидату чи партії сформувати свою програму, лозунги, імідж таким чином, щоб врахувати інтереси, думки, мотивації тих соціальних груп, на які кандидат або партія орієнтуються, на чію підтримку розраховують.

Література:

1. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике : Учеб. пособие для студентов вузов. Изд. 5-е, стер. – М.: Высш. шк., 2000. – 400 с.
2. Кремер Н.Ш. Теория вероятностей и математическая статистика: Учебник для вузов. – 2-е изд., перераб. И доп. – М.: ЮНИТИ-ДАНА, 2004. – 573 с.
3. Паниотто В.И., Максименко В.С. Количественные методы в социологических исследованиях. – Киев, 2003. – 270 с.