ЕКОНОМІЧНА ТА СОЦІАЛЬНА ГЕОГРАФІЯ

УДК 911.3+911.5+910.1

Володимир ГРИЦЕВИЧ

ПЕРСПЕКТИВИ РОЗШИРЕННЯ СФЕРИ ВИКОРИСТАННЯ ЗАКОНУ ЦІПФА В СОЦІАЛЬНО-ЕКОНОМІКО-ГЕОГРАФІЧНИХ ДОСЛІДЖЕННЯХ

Постановка проблеми. "Правило" (або за деякими джерелами "закон", "закономірність", "гіпотеза", "формула") Ціпфа давно використовується в різноманітних наукових дослідженях, зокрема при вивченні суспільно-географічних систем. Найчастіше його застосовують для аналізу чисельності населення в системі міських поселень.

У практиці суспільно-географічних досліджень дуже часто доводиться мати справу з територіальними суспільними системами, що складаються з певних територіальних елементів, і в них є потреба описати розподіл досліджуваної ознаки у розрізі цих елементів, використовуючи закономірності "типу Ціпфа". Однак, досвід таких досліджень показує, що класичний закон Ціпфа у чистому вигляді зустрічається рідко. Досить часто точковий графік, побудований на підставі спостережень з використанням логарифмічних осей, демонструє криволінійну залежність між логарифмами рангу та результуючої ознаки, яку неможливо або недоцільно апроксимувати прямою лінією. Дослідники, що зорієнтовані на класичний закон Ціпфа, змушені взагалі вилучати з розгляду такі випадки.

Аналіз публікацій. Вважається, що вперше закономірність, виражена формулою (1), була використана Г. Ціпфом у 1949 році [5]. Ця закономірність виявилась настільки загальною і вдалою, що вона дуже скоро стала міждисциплінарним інструментом досліджень. У географії суспільства вона прижилась при вивченні розподілу системи міст за чисельністю населення, хоча сфера її застосування може бути набагато ширшою. Методичних робіт, у яких викладається найпростіший варіант правила Ціпфа достатньо багато. Серед них зокрема варто назвати монографії українських економіко-географів [1], [3], [4]. Математико-географічні основи цього методу з прикладом для 100 найбільших міст України (за матеріалами перепису 2001 року) приведені в [2].

Постановка завдання. Вважаю, що криволінійна залежність між логарифмами рангу та результуючої ознаки свідчить про певну специфіку розподілу ознаки і цілком може бути використана в суспільно-географічних дослідженнях при відповідному науково-методичному забезпеченні, мова про яке піде далі.

Нехай досліджувана територіальна система складається з m територіальних елементів будь-якої геометричної природи (точкових, лінійних, або ареальних). Розглянемо ознаку W, яка характеризує елементи і впорядкуємо їх за спаданням значення цієї ознаки. Отже, перший елемент має найбільше значення $W = W_1$ і ранг 1. Останній елемент має найменше значення $W = W_m$ і ранг m. Ставиться задача побудувати математичну модель залежності $W = W(k)$, де k - ранг елемента. Отже, інформаційною базою дослідження тут служить масив спостережень $W_1, W_2, \ldots, W_m$.

Якщо $\ln W$ залежить від $\ln k$ лінійно, тобто $\ln W = a_0 + a_1 \cdot \ln k$, то звідси шляхом потенціювання визначають W у вигляді залежності $W(k) = C \cdot k^{a_1}$, де $C = e^{a_0}$. На практиці завжди $a_1 < 0$, тому традиційно закон Ціпфа записують так

$$W(k) = \frac{C}{k^{b_1}}, \quad (1)$$

де $b_1 = -a_1$ і $b_1 > 0$. Щоправда, такий вигляд не гарантує бажаної інколи рівності $W(1) = W_1$, тому, при особливій потребі її забезпечення, лінійне рівняння в логарифмах записують так $\ln W - \ln W_1 = a_1 \cdot \ln k$ і визначають лише a_1 (або b_1).

Отримані результати. Припустимо, що точковий графік виявив криволінійну залежність $\ln W = f(\ln k)$, де f - деяка нелінійна функція. Зрозуміло, що f можна апроксимувати безліччю способів, але для широкого загалу дослідників потрібно зупинитися на деяких простих і зрозумілих способах. Найприроднішим є перехід від традиційного лінійного рівняння $\ln W = a_0 + a_1 \cdot \ln k$ до квадратичного $\ln W = a_0 + a_1 \cdot \ln k + a_2 \cdot (\ln k)^2$, а далі аналогічно до кубічного рівняння $\ln W = a_0 + a_1 \cdot \ln k + a_2 \cdot (\ln k)^2 + a_3 \cdot (\ln k)^3$ і так далі. Вони дають змогу описати залежність $f(\ln k)$ набагато точніше ніж лінійне рівняння. Додатковою перевагою цих рівнянь є те, що самі коефіцієнти a_0, a_1, a_2, a_3 входять в них лінійно, а тому процедура їх визначення є найпростішою з усіх можливих і полягає в стандартному застосуванні методу найменших квадратів з наступним розвязанням системи лінійних алгебраїчних рівнянь. У цій статті зупинимось на квадратичному та кубічному рівняннях, бо досвід показує, що їх достатньо для переважної більшості випадків.

У квадратичному випадку маємо $\ln W = a_0 + a_1 \cdot \ln k + a_2 \cdot (\ln k)^2$, звідки формально можна визначити $W(k) = \exp\left(a_0 + a_1 \cdot \ln k + a_2 \cdot (\ln k)^2\right) = C \cdot k^{a_1 + a_2 \cdot \ln k}$, де $C = e^{a_0}$. На практиці дуже часто виявляється, що $a_1 < 0$ і $a_2 < 0$, тому квадратична версія закону Ціпфа має вигляд

$$W(k) = \frac{C}{k^{b_1 + b_2 \cdot \ln k}}, \quad (2)$$

де $b_1 = -a_1$, $b_2 = -a_2$ і $b_1 > 0$, $b_2 > 0$.

У кубічному випадку маємо $\ln W = a_0 + a_1 \cdot \ln k + a_2 \cdot (\ln k)^2 + a_3 \cdot (\ln k)^3$, звідки формально визначаємо $W(k) = \exp\left(a_0 + a_1 \cdot \ln k + a_2 \cdot (\ln k)^2 + a_3 \cdot (\ln k)^3\right)$. Така форма запису є дещо громіздкою і мало подібною на первісний закон Ціпфа, тому доцільно її спростити шляхом перетворень. У результаті отримаємо $W(k) = C_0 \cdot (C_2 \cdot k)^{a_1 + a_3 \cdot (\ln k)^2}$, де $C_0 = e^{a_0 - \frac{a_1 \cdot a_2}{a_3}}$, $C_2 = e^{\frac{a_2}{a_3}}$. На практиці дуже часто $a_1 < 0$ і $a_3 < 0$, тому спрощена кубічна версія закону Ціпфа має вигляд

$$W(k) = C_0 \cdot \left(\frac{B_2}{k}\right)^{b_1 + b_3 \cdot (\ln k)^2}, \quad (3)$$

де $B_2 = \dfrac{1}{C_2}$ $b_1 = -a_1$, $b_3 = -a_3$ і $b_1 > 0$, $b_3 > 0$. Для орієнтовного контролю обчислень можна використовувати співвідношення $C_0 \cdot B_2^{b_1} \approx W_1$. Залежно від конкретної ситуації можуть бути й інші спрощені версії кубічного закону Ціпфа.

Розглянемо конкретні приклади застосування лінійного та узагальненого правил Ціпфа у розрізі регіонів України. Статистичні довідники подають інформацію для двадцяти п'яти областей України і окремо для міст Києва та Сімферополя. Досвід досліджень показує, що в одних випадках можна підпорядковувати правилу Ціпфа всі 27 названих суб'єктів, в інших випадках доцільно об'єднувати Сімферополь з Кримом (коли обсяг певного показника в Севастополі непропорційно малий).

Приклади лінійних моделей.

- Кількість суб'єктів ЄДРПОУ на початок 2006 року:

$$W(k) = \frac{195438}{k^{0,7686}},\ k = 1,\ldots,27;$$

якщо поставити умову, $W(1) = W_1$ і включити Севастополь у Крим, то отримаємо модель:

$$W(k) = \frac{183866}{k^{0,735}},\ k = 1,\ldots,26.$$

- Роздрібний товарообіг у 2005 році, млн. грн:

$$W(k) = \frac{18428}{k^{0,8141}},\ k = 1,\ldots,27;$$

при умові, $W(1) = W_1$ і включенні Севастополя до Криму отримується модель:

$$W(k) = \frac{17695,6}{k^{0,7887}},\ k = 1,\ldots,26.$$

- Обсяг виробництва послуг у 2005 році, млн. грн:

$$W(k) = \frac{24441}{k^{1,0965}},\ k = 1,\ldots,27.$$

- Обсяг реалізованих послуг у 2005 році, млн. грн:

$$W(k) = \frac{25591}{k^{1,1148}},\ k = 1,\ldots,27.$$

- Викиди шкідливих речовин в атмосферу в 2005 році, тис. т:

$$W(k) = \frac{2342}{k^{1,2186}},\ k = 1,\ldots,27.$$

Приклади квадратичних моделей (Севастополь включений до Криму).

- Постійне населення на 1 січня 2006 року, осіб:

$$W(k) = \frac{4381931}{k^{0,236+0,0686\cdot\ln k}},\ k = 1,\ldots,26.$$

- Чоловіки на 1 січня 2006 року, осіб:

$$W(k) = \frac{2004690}{k^{0,2257+0,0711\cdot\ln k}},\ k = 1,\ldots,26.$$

- Жінки на 1 січня 2006 року, осіб:

$$W(k) = \frac{2385693}{k^{0,2379+0,0700\cdot\ln k}},\ k = 1,\ldots,26.$$

- Особи, молодші працездатного віку на 1 січня 2006 року, осіб:

$$W(k) = \frac{604405}{k^{0,1775+0,0641\cdot\ln k}},\ k = 1,\ldots,26.$$

- Особи працездатного віку на 1 січня 2006 року, осіб:

$$W(k) = \frac{2711466}{k^{0,2230+0,0793\cdot\ln k}},\ k = 1,\ldots,26.$$

- ВІЛ-інфіковані у 2005 році, осіб:

$$W(k)=\frac{2607}{k^{0,1546+0,2908\cdot\ln k}},\ k=1,\ldots,26.$$

- Хворі на СНІД у 2005 році, осіб:

$$W(k)=\frac{1139}{k^{0,3738+0,3126\cdot\ln k}},\ k=1,\ldots,26.$$

Приклади кубічних моделей (Севастополь включений до Криму).

- Особи, старші працездатного віку на 1 січня 2006 року, осіб:

$$W(k)=42579\cdot\left(\frac{65,83}{k}\right)^{0,7967+0,0839\cdot(\ln k)^2},\ k=1,\ldots,26$$

- Вантажообіг автомобільного транспорту в 2005 році, млн. ткм:

$$W(k)=448,9\cdot\left(\frac{25,97}{k}\right)^{0,6498+0,1137\cdot(\ln k)^2},\ k=1,\ldots,26.$$

- Імпорт у 2005 році, млн. дол. США:

$$W(k)=0,3839\cdot\left(\frac{83,93}{k}\right)^{2,3361+0,2263\cdot(\ln k)^2},\ k=1,\ldots,26.$$

- Обсяг роздрібної торгівлі в 2005 році, млн. грн:

$$W(k)=4,1613\cdot\left(\frac{4,711}{k}\right)^{0,2291+0,0371\cdot(\ln k)^2},\ k=1,\ldots,26.$$

- Кількість малих підприємств у 2005 році, одиниць:

$$W(k)=802,2\cdot\left(\frac{79,3}{k}\right)^{0,9459+0,0509\cdot(\ln k)^2},\ k=1,\ldots,26.$$

- Кількість інвалідів у 2005 році, осіб:

$$W(k)=18738\cdot\left(\frac{45,18}{k}\right)^{0,6641+0,0877\cdot(\ln k)^2},\ k=1,\ldots,26.$$

- Кількість зареєстрованих злочинів у 2005 році, тис. випадків:

$$W(k)=55,66\cdot\left(\frac{1,1212}{k}\right)^{0,3012+0,0367\cdot(\ln k)^2},\ k=1,\ldots,26.$$

Головним параметром, який у першу чергу визначає характер залежності $W=W(k)$ є коефіцієнт при найвищому степені логарифма. У лінійному рівнянні це a_1, у квадратичному - a_2, а в кубічному a_3. Близькі значення головного параметра для різних показників є хорошою евристичною ознакою і можуть свідчити про подібність їх просторового розподілу, а далі про закономірності або навіть закони їхньої територіальної відповідності. Зокрема, для наведених вище кубічних моделей звертає на себе увагу близькість параметра a_3 (або b_3) для таких пар показників:

1. кількість осіб старших працездатного віку і кількість інвалідів;
2. обсяг роздрібної торгівлі і кількість зареєстрованих злочинів.

Розглянемо детальніше останню пару (рис. 1), бо її зміст є достатньо резонансним для суспільства. Наведений точковий графік чітко показує наявність залежності між показниками. Цю залежність (у поданих одиницях вимірювання) можна виразити трендовим рівнянням $y = 3{,}77 \cdot e^{-0{,}47 \cdot x}$.

Висновки. Із проведених досліджень випливає наступне:

1. Закономірності типу "Правила Ціпфа" є ефективним інструментом суспільно-географічних досліджень.
2. Визначення закономірностей типу Ціпфа є одним із методів математичної географії суспільства.

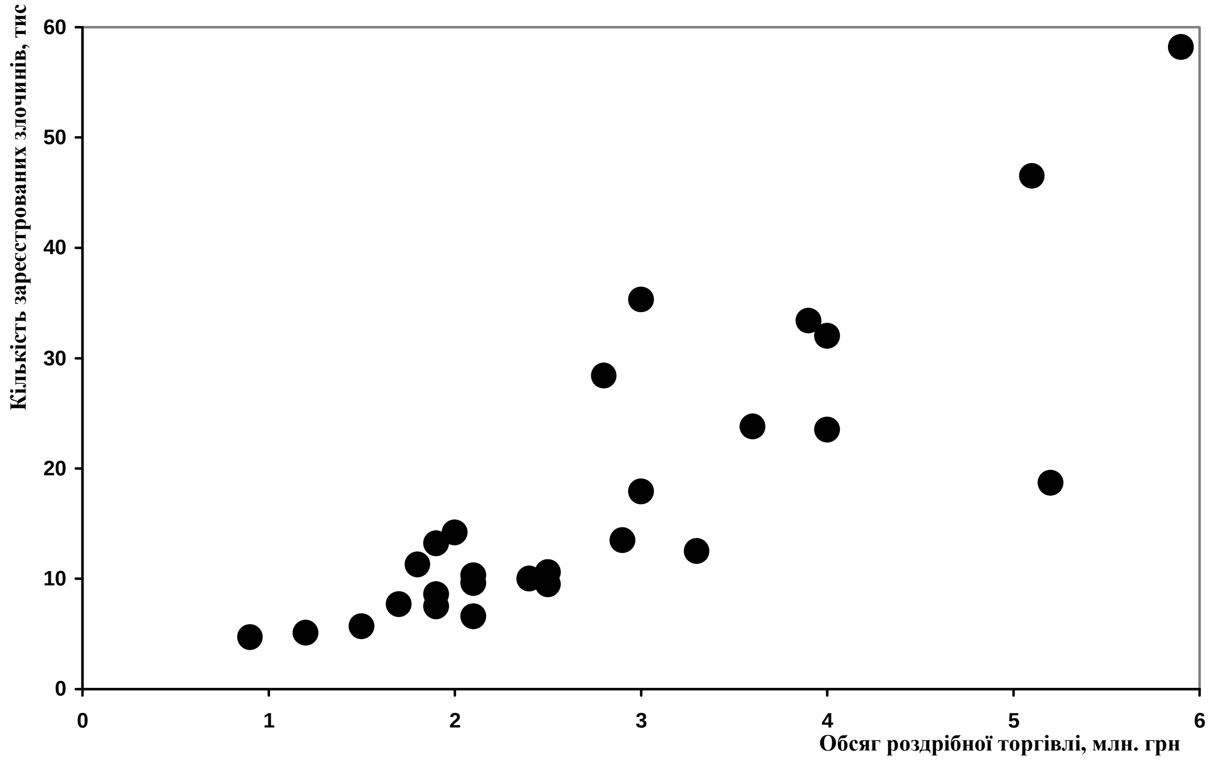

Рис. 1. Залежність між обсягом роздрібної торгівлі та кількістю зареєстрованих злочинів у розрізі регіонів України в 2005 році.

3. Квадратичний і, особливо, кубічний, варіанти правила Ціпфа дають змогу набагато точніше описати реальний розподіл показника, ніж традиційний лінійний варіант.
4. Близькість головних параметрів рівняння, що описує закономірність типу Ціпфа, для різних показників допомагає виявити територіальну відповідність між явищами, які описуються цими показниками.

Література:

1. *Голиков А.П., Черванев И.Г. Трофимов А.М.* Математические методы в географии. –Харків: Вища школа, 1986.
2. *Грицевич В.С.* Математичні методи в демогеографії. Текст лекції. -Львів: Видавничий центр ЛНУ імені Івана Франка, 2003. –30 с.
3. *Топчієв О.Г.* Суспільно-географічні дослідження: методологія, методи, методики. –Одеса: Астропринт, 2005.
4. *Шаблій О.І.* Математичні методи в соціально-економічній географії. –Львів: Світ, 1994.
5. *Zipf G.K.* Human Behaviour and the Principle of Least-Effort. -Cambridge MA: Addison-Wesley, 1949.

Summary:

V. Grytsevych. THE PROSPECTS OF EXPANSION OF SPHERE OF USE ZIPF'S RULE IN SOCIAL ECONOMICAL AND GEOGRAPHICAL RESEARCHES.

The generalization of Zipf's model is offered for cases, when parameters of social-geographical phenomena are characterized by nonlinear dependence on their grades in a logarithmic scale. Developed concrete method for construction the generalized models of "Zipf's type" for quadratic and cube nonlinear dependence. Shown numerous examples of such mathematical-geographical modeling in regions of Ukraine.