ПРОБЛЕМИ ЗАСТОСУВАННЯ ЙМОВІРНІСНИХ МЕТОДІВ В ЕКОНОМІЦІ

Джунь Й.В.
професор, доктор фізико-математичних наук
Рівненського економіко-
туманітарного інституту

Вражаючі події останніх років XX століття, а саме - фінансова криза в Мексіці, Таїланді, Японії, Південній Кореї, Індонезії з усією очевидністю підтверджують існування істотної іймовірнісної компоненти у розвитку економічних процесів.

Помилково вважати, що створенням нових баз даних чи шляхом збільшення обсягів економічної інформації можна захистити ці процеси від ударів імовірної стихії. Цьому протидіють такі потужні фактори:

- наявність тіньових процесів в економіці, які майже не відображаються або повністю проходять за межами офіційної статистики, причому, тіньові процеси - це реалії як розвинених, так і відсталих в економічному розвитку країн;

- по-друге, припинення холодної війни зовсім не означає припинення війни інформаційної, а, скоріше, її зростання, оскільки дезінформація є старим, але невмирущим засобом захисту економічних інтересів, (тут слушно згадати знаменитий вислів Б. Дізраелі, який досить точно відображає суть інформаційної війни: "Є три види брехні: брехня, брехня нагла і статистика."[1]);

- по-третє, і це головне, XX століття є епохою небаченої економічної боротьби держав, корпорацій, фірм, діяльність яких невидимими фінансовими потоками живить корупцію, лоббізм, які поступово, крок за кроком здатні перетворити в руїну економіку цілих держав. Неефективні, кровопролитні, занадто дорогі збройні конфлікти якось непомітно переросли в війни економічні - непомітні, досконало забезпечені дезінформацією і ефективні.

Окреслені вище фактори, які неперервно взаємодіють, надають економіці деяких держав, а, в цілому, і світовій економіці досить некерований характер. Тому однією з основних задач імовірних методів аналізу економічної інформації, а точніше - однією з проблем економічної науки, є визначення того імовірного ряду держав, які можуть стати епіцентрами майбутніх фінансових потрясінь. Другою, не менш важливою проблемою, є прорахунки імовірності того, чи є названі фінансові потрясіння вісниками більш страшних - світових потрясінь.

Підсумовуючи ці висновки, потрібно сказати, що названі три причини інформаційної недостатності банків даних обумовлюють значення імовірнисних процедур аналізу економічної інформації. А тому важливо окреслити не лише інформаційні, а й сучасні аспекти імовірнісного аналізу даних, і ті його проблеми, які ще не зовсім усвідомлюються економістами. Це, перш за все, математичні проблеми аналізу даних, які виникли порівняно недавно, в останні 20-30 років, після того, як класичний теорегико-ймовірнісний апарат статистичного аналізу масових спостережень був уже розроблений. Фундаментом цього апарату є, по суті, два математичних постулати:

1. Це так званий "закон великих чисел ", який є основою теорії спроможних оцінок.

2. Постулат про нормальний розподіл масових спостережень об'єкта, які виконані в статистично однорідних умовах.

Останній постулат є основою таких класичних процедур, як x^2 - критерій, t-критерій, F-критерій чи теорії кореляції.

Питання, яке ми досліджуємо, полягає в наступному: підтверджують чи не підтверджують ряди сучасних спостережень найвищої якості спроможність цих двох постулатів? Такими статистичними рядами найвищої якості є ряди похибок спостережень фундаментальних сталих у фізиці, які наука вважає практично незмінними. Для ілюстрації ми наводимо лише частину отриманих результатів, які є досить показовими і цікавими.

На рис.1. дана діаграма, яка взята з роботи [2]. Діаграма показує у відносних одиницях, як з часом змінювались значення деяких фундаментальних констант, а також оцінки відповідних стандартних відхилень. Автор роботи [2] відмічає незвичний факт: різниця між значеннями, отриманими в різні роки, істотно перевищує величину випадкових помилок. Інакше кажучи, для цілого ряду попередніх результатів систематична помилка істотно перевищує випадкову.

Подібна ситуація описана в роботі американських вчених [3]. Процитуємо яскравий висновок цієї роботи: "Результати, отримані двома лабораторіями після кількох років роботи з застосуванням сучасного обладнання і використанням великої кількості докторів наук, відрізнялись на цілих 25%, але заявлені похибки були в 10 раз меншими".

Ще більш вражаючі результати досліджень наведені в роботі [41 для похибок відтворення 17 офіційно атестованих значень різних фізичних величин у 8 еталонах. Обробка цих даних показала, що відношення невиключеної систематичної помилки а до випадкової 8 результату виміру має такий розподіл:

г а
1 ь — в 12 % випадків;
о

m/e S
і — 2 в 25% випадків;
о

Висновок є таким: всі наведені вище результати, отримані для статистичних рядів найвищої якості, не підтверджують закону великих чисел. Правда полягає в тому, що. як це сформулював П. Е. Ельясберг в роботі [5]: „... в дійсності властивість спроможності оцінок ніколи не здійснюється на практиці і, починаючи з деякого моменту, подальше збільшення обсягів інформації' шляхом спостережень не збільшує точності результату

Чи можуть економісти не звертати уваги на ці нові результати, отримані в фундаментальних науках - науках, які завжди йшли попереду науково-

технічного прогресу. Є надія, що ні. І, найперше, тому, що ці результати прямо показують на нове небезпечне джерело помилок, обумовлених невідповідністю першого математичного постулата реальній практиці спостережень.

Розглянемо тепер питання про спроможність другого математичного постулата із зазначених нами. З цією метою ми наводимо цікавий рис. 2, на якому в площині А о Е вказані області всіх типів кривих розподілів, які розроблені статистичною школою Неймана - Пірсона: *S* - ексцес, А - асиметрія розподілу.

По осі п відкладено обсяги проаналізованих статистичних рядів дуже високої якості: фізичних, гравіметричних, астрономічних. З рис.2 бачимо, що точка N. яка відповідає на площині А о Е нормальному розподілу, знаходиться далеко на периферії реальних однорідних статистичних рядів високої якості. Дійсні розподіли похибок зміщені від точки N в область симетричного розподілу Пірсона VII типу, який має таку щільність:

$$y = \frac{\Gamma(m+1)}{\sqrt{2\pi\left(m-\frac{1}{2}\right)}\Gamma\left(m+\frac{1}{2}\right)\sigma}\cdot\left[1+\frac{1}{2M}\left(\frac{x-a}{\sigma}\right)^2\right]^{-m}, \quad (1)$$

де α, σ, m - параметри розподілу (1),

$M = \left(m-\frac{1}{2}\right)/m^2$, $\Gamma(m+1)$, $\Gamma\left(m+\frac{1}{2}\right)$ - гамма функція

Зазначимо, що розподіл Пірсона VII типу є узагальненням форми розподілу Стьюдента. Параметр m може розглядатись як міра відхилення розподілу (1) від нормального, і він пов'язаний таким співвідношенням з числом ступенів вільності у розподілу Стьюдента:

$$«=^2\text{-j-,m)l,5} \quad (2)$$

В наш час статистики - математики дискутують про те, яке значення m можна застосувати для характеристики нового загального закону похибок. В роботі [б] підкреслено, що дійсно незалежні ряди похибок за даними Г. Джефріса з у=3,5-4,5 , а П. Хьюберг (усне повідомлення) для незалежних статистичних спостережень високої якості назвав, як найкращий приклад, розподіл , тобто t-розподіл Стьк&ента з у=3. Тобто, якщо з результатів статистичних спостережень повністю виключити систематичні похибки, то випадкові помилки матимуть не нормальний розподіл, а істотно будуть відхилятись від нього в області розподілу Стьюдента зу=3. Напевне, ще рано говорити про те, що в нашому фізичному світі принципово не можуть бути реалізовані умови центральної граничної теореми теорії ймовірностей, але можна вже говорити про те, що ці умови дуже далекі від фізичної реальності і практично ніколи не можуть бути реалізовані.

Якщо в дійсності випадкові похибки мають /3 -розподіл, то m=(v+l)/2=2. З цього випливає, що середнє арифметичне вже не є ефективною оцінкою математичного сподівання. Для побудови такої оцінки потрібно визначити вагову' функцію [7] розподілу

$$W_i = \frac{y'}{y\cdot V_i}, \quad (3)$$

де $_i = x_i - a$, а - математичне сподівання.

Зазначимо, що в класичних методах обробки статистичних даних вагова функція (3) є сталою для всіх , (нормальний розподіл). Для розподілу Пірсона VII типу маємо:

$$Wi = \left[\frac{m-1.5}{m}\cdot M_2 + \frac{V_i^2}{2m}\right]^{-1} \quad (4)$$

Нехай M_2=1, m=2, тоді маємо:

$$W_i = \frac{4}{1+V_i^2} \quad (5)$$

При ν_i=0 (для спостережень в околі а) W_i=4.

При ν_i=3, W_3=0,4, тобто, відношення $\frac{W_0}{W_3} = 10$ (6)

Така відмінність ваг окремих спостережень статистичного ряду є вже істотною, наприклад, при застосуванні середніх показників чи при спектральному аналізі рядів динаміки.

Із всього сказаного вище можна зробити такий важливий висновок: застосування класичних процедур при обробці економічної інформації, незважаючи на їх поважний математичний вигляд, має лише приблизний, а не науково обгрунтований характер через те, що фундаментальна гіпотеза нормальності, покладена в основу цих критеріїв, як правило, не реалізується на практиці. Підтвердження гіпотези нормальності статистичних рядів на основі X^2-критерію для спостережень обсягом менше 300 обумовлене лише недостатнім обсягом інформації. Статистичні ряди обсягом 500 спостережень і більше мають вже явні ознаки ненормальності. Цієї надзвичайно важливої в теорії обставини якраз і не враховують класичні методи математичної статистики, приписуючи всім розподілам похибок ексцес є-0, незважаючи на те, що реальні ексцеси випадкових похибок спостережень істотно більші нуля. Зазначимо також, що метою нашого дослідження не є критика існуючого стану застосування методу аналізу' даних; ми лише намагались висвітлити суть сучасних інформаційних і математичних проблем, пов'язаних із застосуванням імовірнісно - статистичних методів аналізу великих масивів економічної інформації.

Література

1. G.A. Kimble. Haw to use (and misuse) Statistics. Englewood Cliffs, N. I. Prentice-Hall, inc., 1978
2. А Н. Зайдель. Погрешности измерений физических величин. Л.: Наука, 1985.
3. National Standard Reference Data System. Natl. Bur. Stand. Techn. News. Bull, 1969, oktober, p.230.
4. В. В. Василенко. Метрологическая служба в СССР. Всес. НИИ технической информации, классификации и кодирования. 1988, вып. 9,с. 1.
5. П. Е. Эльясберг. Измерительная информация: сколько её нужно? Как обрабатывать. М.: Наука. 1983.
6. Ф. Хампель, Э. Рониетги, П. Рауссеу, В. Штаэль. Робастность в статистике. Пер. с англ. М.: Мир, 1989.
7. H. Jeffreys. Theory of probability. Oxford, 3-d. ed. Clarendon Press, 1961.

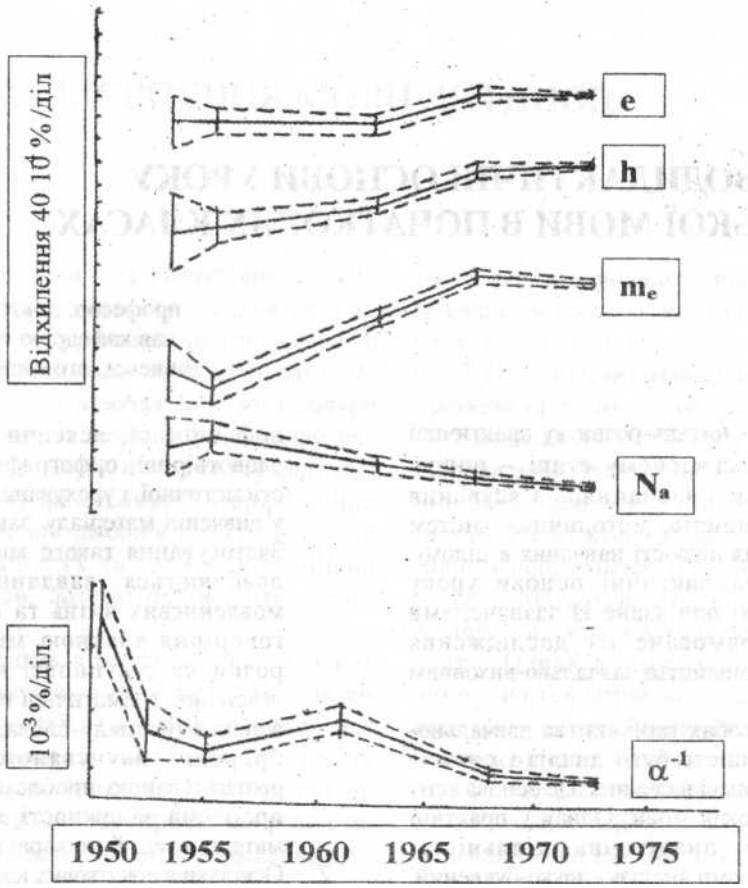

Рис. 1. Зміна відносних значень деяких фундаментальних сталих і відносних стандартних похибок, їх оцінок: e - заряду електрона; її - стала Планка; $т_e$ - маси електрона; N_A - число Авогадро; $ОС^1$ - стала надтонкої структури.

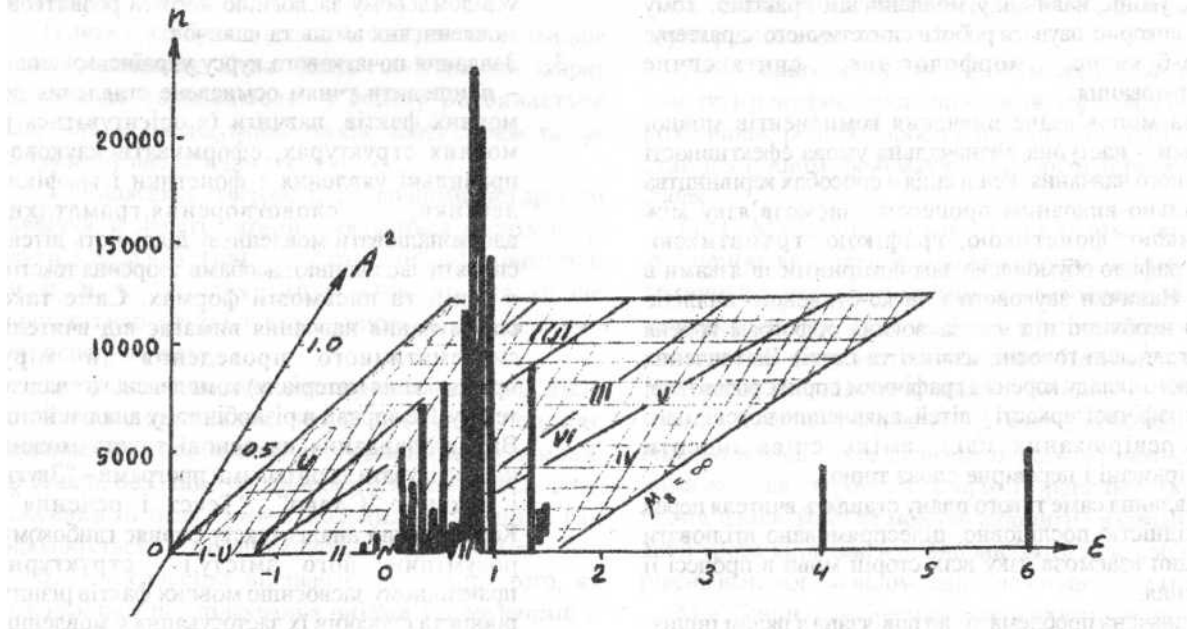

Рис.2. Розміщення статистичних розподілів похибок спостережень високої якості по областях імовірнісних сімейств Пірсона.