УДК 330.1317:519.81:65.016

Використання коефіцієнтів асиметрії та ексцесу у параметричних статистичних моделях

Василь Єлейко[1], Євген Пенцак[2]

[1] д. е. н., професор, Львівська комерційна академія, вул. Туган-Барановського, 10, Львів, 79010
[2] к. ф.-м. н., Львівський інститут менеджменту, вул. Ліська, 16, Львів, 79015, e-mail: YP@fame.ch

Пропонується методика використання коефіцієнтів асиметрії й ексцесу параметричних розподілів випадкових величин у моделях фінансової економетрики й економетрики охорони здоров'я, для яких є характерними наявність «товстих хвостів» і великого значення коефіцієнта асиметрії. Візуалізація нормалізованих моментів третього та четвертого порядків емпіричного розподілу дозволяє адекватно вибрати параметричну сім'ю розподілів, відповідні нормалізовані моменти якої покривають їх емпіричні аналоги. На прикладах показано, що стандартні параметричні функції щільностей, які використовуються для моделювання асиметрії й ексцесу, не можуть генерувати емпіричних нормалізованих центральних моментів третього та четвертого порядків для даних витрат на лікування та прибутковості активів на українській фондовій біржі.

Ключові слова: коефіцієнт асиметрії, коефіцієнт ексцесу, ненормально розподілена випадкова величина, функція щільності розподілу.

Вступ. Відомо, що багато залежних змінних, наприклад, витрати на лікування, дохід, заробітна плата, прибутковість активів, володіють специфічними рисами, які суперечать припущенню про нормальний розподіл. Зокрема, розподіл витрат на лікування типово демонструє надмірну асиметрію (*skewness*) і надлишковий ексцес (*kurtosis*), що навіть звичний для моделювання цього явища логнормальний розподіл не може врахувати особливостей емпіричного розподілу витрат [7].

Відомо багато статистичних та економетричних методів, які дозволяють моделювати зазначені вище специфічні риси цих розподілів. Серед цих методів розрізняють непараметричні методи оцінювання, метод моментів і різні робастні методики, які є стійкими до незначних трансформацій баз даних. У даній роботі зупинимося на параметричних методах оцінки економетричних моделей, у яких збурення не є нормально розподіленими. Емпіричні дослідження показують, що навіть дуже добра апроксимація емпіричного розподілу в його «голові» не гарантує прийнятних властивостей для відповідних прогнозів. Тому фахівці в галузі фінансової економетрики й економетрики у сфері охорони здоров'я звертають особливу увагу на здатність параметричного розподілу апроксимувати нормалізовані моменти емпіричного розподілу [2].

У роботі зроблено огляд основних розподілів, які пропонуються для моделювання ненормально розподілених випадкових величин та визначено їхні теоретичні області покриття нормалізованих моментів третього (коефіцієнт асиметрії) та четвертого порядків (коефіцієнт ексцесу). Для дослідників ця інформація є важливим індикатором щодо придатності обраного розподілу для моделювання досліджуваного явища. У роботі наведено приклади застосування запропонованої методики до аналізу витрат на лікування серцево-судинних захворювань на основі даних швейцарського госпіталю CHUV, а також аналізу даних прибутковості компаній із найбільшою капіталізацією на фондовому ринку України (ПФТС). Отримані результати можуть бути використані також для оцінки вартості під ризиком (VaR) у ризик-менеджменті.

У роботі наведено стандартні параметричні функції щільності, які використовуються для моделювання асиметрії та «товстих хвостів» емпіричних розподілів; описано проблему моментів Ґамбургера та визначено теоретичну границю покриття коефіцієнтами асиметрії й ексцесу. Проілюстровано нормалізовані моменти третього та четвертого порядків для описаних вище стандартних розподілів, закономірності розподілів емпіричних коефіцієнтів ексцесу й асиметрії для прибутковості «голубих фішок» української фондової біржі, а також витрат на лікування різноманітних серцево-судинних захворювань на основі даних швейцарського госпіталю CHUV.

1. Стандартні функції щільності, які використовуються для моделювання асиметрії та «товстих хвостів»

Наведемо типові приклади функцій щільності, включені у стандартні статистичні й економетричні пакети, які допускають оцінку їх параметрів методом максимальної правдоподібності вбудованими у них засобами. Оскільки нас будуть цікавити теоретичні області покриття розподілами нормалізованих третіх і четвертих моментів, то запишемо також формули для їх обчислення. Для кількісної оцінки відхилення від нормального розподілу дослідники використовують міри асиметрії й ексцесу [5], що визначаються як

$$\gamma_1 = \frac{m_3}{m_2^{3/2}}, \quad \beta_2 = \frac{m_4}{m_2^{2}}$$

відповідно, де m_i — центральний момент порядку i. Коефіцієнт γ_1, відомий як «стандартизований третій центральний момент», вказує на відхилення від симетрії і називається коефіцієнтом асиметрії. Коли розподіл є симетричним, то $\gamma_1 = 0$. Проте, можна отримати $\gamma_1 = 0$ і тоді, коли розподіл не є цілком симетричним. Коефіцієнт $\gamma_2 = \beta_2 - 3$ (β_2 відомий як «стандартизований четвертий центральний момент») допомагає класифікувати розподіли відносно їх поведінки у «хвостах» і називається коефіцієнтом ексцесу.

Для повноти викладу розпочнемо з функції щільності нормального розподілу.

- Функція щільності для *нормального розподілу*

$$f_N(x;\mu,\sigma) = \frac{1}{2\sqrt{\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \tag{1}$$

де $\mu \in (-\infty;\infty)$ і $\sigma \in (0;\infty)$. Для нормального розподілу $\gamma_1 = 0$ і $\gamma_2 = 0$.

- Функція щільності для *логнормального розподілу*

$$f_{LN}(x;\mu,\sigma) = \frac{1}{2x\sqrt{\pi}\sigma}\exp\left[-\frac{(\ln x-\mu)^2}{2\sigma^2}\right], \tag{2}$$

де $\mu \in (-\infty;\infty)$ і $\sigma \in (0;\infty)$. Відомо, що для логнормального розподілу

$$\gamma_1 = \frac{\exp(3\sigma^2) - 3\exp(\sigma^2) + 2}{\left[\exp(\sigma^2) - 1\right]^{3/2}},$$

$$\gamma_2 = \frac{\exp(6\sigma^2) - 4\exp(3\sigma^2) - 3\exp(2\sigma^2) + 12\exp(\sigma^2) - 6}{\left[\exp(\sigma^2) - 1\right]^2}.$$

- Функція щільності для *гама розподілу*

$$f_G(x;a,b) = \frac{x^{a-1}}{b^a}\exp(x/b), \quad x \in (0;\infty), \tag{3}$$

де $a \in (0;\infty)$ і $b \in (0;\infty)$. Відомо, що для гама розподілу

$$\gamma_1 = \frac{\Gamma(3+a) - 3a\Gamma(2+a) + 2a^3\Gamma(a)}{a^{3/2}\Gamma(a)},$$

$$\gamma_2 = \frac{\Gamma(4+a) - 4a\Gamma(3+a) + 6a^2\Gamma(2+a) - 3a^4\Gamma(a) - 3a^2\Gamma(a)}{a^2\Gamma(a)}.$$

- Функція щільності для *розподілу Вейбула*

$$f_W(x;a,b) = \frac{bx^{b-1}}{a^b}\exp\left[-\left(\frac{x}{a}\right)^b\right], \tag{4}$$

де $a \in (0;\infty)$ і $b \in (0;\infty)$. Відомо, що для розподілу Вейбула

$$\gamma_1 = \frac{\Gamma(1+3/b) - 3\Gamma(1+2/b)\Gamma(1+1/b) + 2\Gamma^3(1+1/b)}{\left[\Gamma(1+2/b) - \Gamma^2(1+1/b)\right]^{3/2}},$$

$$\gamma_2 = \frac{\Gamma(1+4/b) - 4\Gamma(1+3/b)\Gamma(1+1/b) + 12\Gamma(1+2/b)\Gamma^2(1+1/b)}{\left[\Gamma(1+2/b) - \Gamma^2(1+1/b)\right]^2} -$$

$$-\frac{3\Gamma^2(1+2/b) - 6\Gamma^4(1+1/b)}{\left[\Gamma(1+2/b) - \Gamma^2(1+1/b)\right]^2}.$$

- Функція щільності для *розподілу Стьюдента*

$$f_{ST}(x;\nu) = \frac{\Gamma\left[(\nu+1)/2\right]}{\Gamma(\nu/2)}\frac{1}{\sqrt{\nu\pi}}\frac{1}{(1+x^2/\nu)^{(\nu+1)/2}}, \tag{5}$$

де $\nu \in (0;\infty)$. Відомо, що для розподілу Стьюдента

$$\gamma_1 = 0\,, \qquad \gamma_2 = \frac{6}{\nu - 4}.$$

- Функція щільності для *розподілу Рейлі*

$$f_R(x;\sigma) = \frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right), \quad x \in (0;\infty), \tag{6}$$

де $\sigma \in (0;\infty)$. Відомо, що для розподілу Рейлі

$$\gamma_1 = \frac{2\sqrt{\pi}(\pi-3)}{(4-\pi)^{3/2}}, \qquad \gamma_2 = \frac{(3\pi-8)^2}{(4-\pi)^2}.$$

- Функція щільності для *оберненого розподілу Гауса*

$$f_{IG}(x;a,b) = \frac{\sqrt{b}}{\exp[0{,}5\ln(x/a)]} f_N\left[\frac{\sqrt{b}(x/a-1)}{\exp[0{,}5\ln(x/a)]};0{,}1\right], \quad x \in (0;\infty), \tag{7}$$

де $a \in (0;\infty)$ і $b \in (0;\infty)$. Відомо, що для оберненого розподілу Гауса

$$\gamma_1 = 3\big/\sqrt{b}\,, \qquad \gamma_2 = 15/b.$$

- Функція щільності для *розподілу Бірнбаума-Саундерса*

$$f_{BS}(x;a,b) = \frac{\sqrt{x/a}+\sqrt{a/x}}{2bx} f_N\left(\frac{\sqrt{x/a}-\sqrt{a/x}}{b};0{,}1\right), \quad x \in (0;\infty), \tag{8}$$

де $a \in (0;\infty)$ і $b \in (0;\infty)$. Відомо, що для розподілу Бірнбаума-Саундерса

$$\gamma_1 = \frac{16a^2(11a^2+6)}{(5a^2+4)^3}, \qquad \gamma_2 = \frac{6a^2(93a^2+41)}{(5a^2+4)^2}.$$

- Функція щільності для *бета розподілу*

$$f_B(x;a,b) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, \quad x \in (0;1), \tag{9}$$

де $a \in (0;\infty)$ і $b \in (0;\infty)$. Відомо, що для бета розподілу

$$\gamma_1 = \frac{2(b-a)}{(a+b+2)}\sqrt{\frac{a+b+1}{ab}}, \qquad \gamma_2 = 6\frac{a^3 - a^2(2b-1) + b^2(b+1) - 2ab(b+2)}{ab(a+b+2)(a+b+3)}.$$

- Функція щільності для *розподілу Фішера*

$$f_F(x;\nu) = \frac{\Gamma((\nu_1+\nu_2)/2)}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)}\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} \frac{x^{\nu_1/2-1}}{\left[1+(\nu_1/\nu_2)x^2\right]^{(\nu_1+\nu_2)/2}}, \tag{10}$$

де $\nu_1 \in (0;\infty)$ і $\nu_2 \in (0;\infty)$. Відомо, що для розподілу Фішера

$$\gamma_1 = \frac{2\nu_1+\nu_2-2}{\nu_2-6}\sqrt{\frac{8(\nu_2-4)}{\nu_1(\nu_1+\nu_2-2)}},$$

$$\gamma_2 = 12\frac{20\nu_2 - 8\nu_2^2 + \nu_2^3 + 44\nu_1 - 32\nu_1\nu_2 + 5\nu_1\nu_2^2 - 22\nu_1^2 + 5\nu_1^2\nu_2 - 16}{\nu_1(\nu_2-6)(\nu_2-8)(\nu_1+\nu_2-2)} - 3.$$

Бачимо, що для розподілу Рейлі та нормального розподілу значення коефіцієнтів асиметрії й ексцесу є сталими, а для розподілу Стьюдента значення коефіцієнта асиметрії дорівнює нулеві. Тому зупинимось лише на геометричній ілюстрації відповідних коефіцієнтів для решти розподілів.

2. Проблема моментів

Засади характеризації можливого покриття коефіцієнтами асиметрії та ексцесу ґрунтуються на теоретичних дослідженнях Гамбургера [2] та стосуються проблеми існування такої неспадної функції α для послідовності нецентральних моментів μ_j, $j = \overline{1, M}$, що

$$\mu_j = \int_{-\infty}^{\infty} z^j d\alpha(z).$$

У роботі [9] було доведено, що достатньою умовою для цього є додатня визначеність послідовності μ_j, тобто

$$\mu_0 \geq 0, \quad \begin{vmatrix} \mu_0 & \mu_1 \\ \mu_1 & \mu_2 \end{vmatrix} \geq 0, \quad \begin{vmatrix} \mu_0 & \mu_1 & \mu_2 \\ \mu_1 & \mu_2 & \mu_3 \\ \mu_2 & \mu_3 & \mu_4 \end{vmatrix} \geq 0, \ \ldots \tag{11}$$

Зокрема, для проблеми чотирьох моментів з умови додатньої визначеності (11) отримаємо таке співвідношення між коефіцієнтом асиметрії й ексцесу

$$\gamma_1^2 < \beta_2 - 1. \tag{12}$$

А тому, для даного значення ексцесу може бути зреалізована лише обмежена область значень коефіцієнта асиметрії [2]. Проте, область покриття для коефіцієнтів асиметрії й ексцесу може бути навіть значно вужчою від заданої теоретичним обмеженням (12). Ґрунтуючись на моментах емпіричного розподілу, зокрема, на коефіцієнтах асиметрії й ексцесу, деякі дослідники намагаються будувати ентропійні щільності [4], щільності Ґрема-Чарлі [3] чи знайти наближення з допомогою параметричних сімей щільностей розподілів. Не аналізуючи недоліки перших двох способів, зосередимося на останньому підході.

3. Графічна ілюстрація коефіцієнтів асиметрії й ексцесу для стандартних розподілів. Застосування

Наведемо графічну ілюстрацію (рис. 1) коефіцієнта асиметрії (вісь y) й ексцесу (вісь x) для деяких параметричних сімей розподілів, які розглядалися вище. Із рис. 2 бачимо, що розподіл Фішера є більш гнучким у плані генерування області покриття коефіцієнтів асиметрії й ексцесу порівняно з іншими стандартними розподілами, які показані на рис. 1.

Багато залежних змінних, наприклад, витрати на лікування, дохід, заробітна плата, прибутковість активів, не є нормально розподіленими. До того ж, часто витрати

ISSN 1816-1545 Фізико-математичне моделювання та інформаційні технології
2006, вип. 4, 114-122

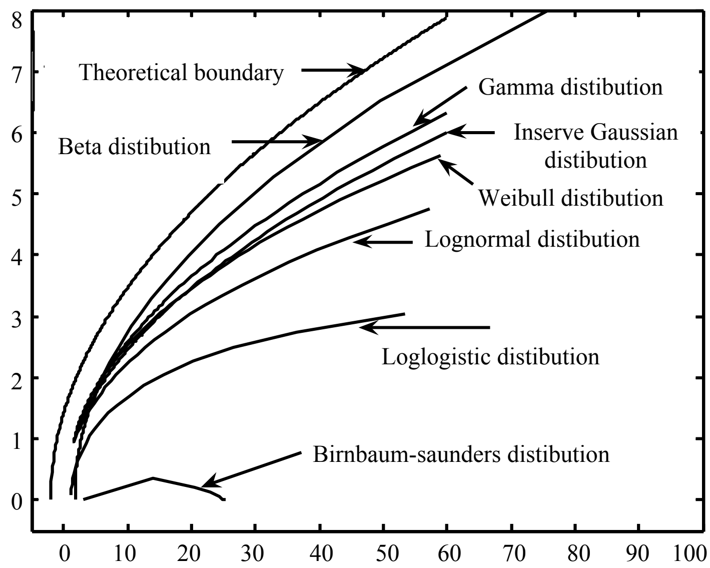

Рис. 1. Коефіцієнти ексцесу й асиметрії
для деяких параметричних розподілів

на лікування характеризуються надмірною асиметрією та «товстими хвостами». Розлянемо дві бази даних: витрати на лікування серцево-судинної системи у швейцарському госпіталі CHUV і прибутковість акцій на українській фондовій біржі (ПФТС).

3.1. Витрати на лікування серцево-судинної системи у швейцарському госпіталі CHUV. Розглянемо емпіричні значення коефіцієнтів асиметрії й ексцесу для різних категорій пацієнтів, які пройшли лікування серцево-судинної системи

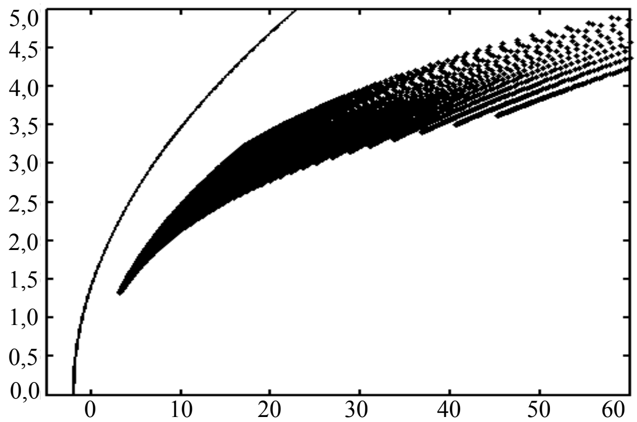

Рис. 2. Графічна ілюстрація коефіцієнтів ексцесу та асиметрії
для розподілу Фішера, *F*

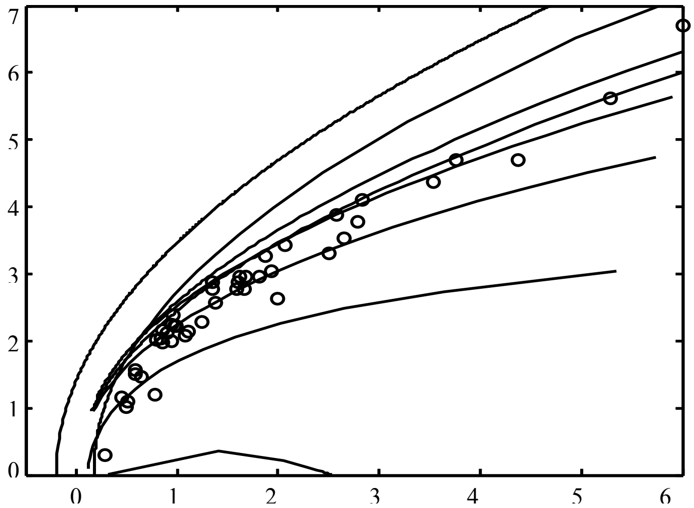

Рис. 3. Емпіричні значення коефіцієнтів ексцесу та асиметрії витрат на лікування серцево-судинних захворювань різних категорій пацієнтів у швейцарському госпіталі CHUV та їх теоретичні аналоги для різних параметричних сімей функцій щільності

у швейцарському госпіталі CHUV. Для графічної ілюстрації (рис. 3) співставимо ці емпіричні характеристики розподілу витрат на лікування із фрагментами кривих, які описують теоретичні значення коефіцієнтів асиметрії й ексцесу. Аналогічно зобразимо емпіричні значення коефіцієнтів асиметрії й ексцесу порівняно з областю покриття відповідних значень розподілом Фішера (рис. 4).

Провівши візуальний аналіз, бачимо, що жоден із запропонованих розподілів не покриває область емпіричних значень коефіцієнтів асиметрії й ексцесу для бази даних витрат на лікування серцево-судинної системи.

3.2. Прибутковість акцій на українській фондовій біржі (ПФТС). Запишемо результати обчислень коефіцієнтів асиметрії й ексцесу для прибутковості «голубих фішок» української фондової біржі.

Таблиця 1

Коефіцієнти асиметрії й ексцесу прибутковості українських «голубих фішок»

Компанія	Центр-енерго	Дніпро-енерго	Захід-енерго	Київ-енерго	Укр-нафта	Стірол	Укр-телеком	Донбас-енерго
коефіцієнт асиметрії	0,05	– 0,51	0,06	– 0,56	– 1,21	0,51	0,10	– 0,39
коефіцієнт ексцесу	– 2,30	1,21	– 0,19	7,28	10,96	– 2,24	– 0,17	0,24

Можна показати, що жодна із запропонованих сімей функцій щільності не покриває область отриманих коефіцієнтів асиметрії й ексцесу. Таким чином, фінансове моделювання вимагає ширшого спектру параметричних сімей функцій щільності, до яких належать різноманітні узагальнення гама розподілу, розподілів

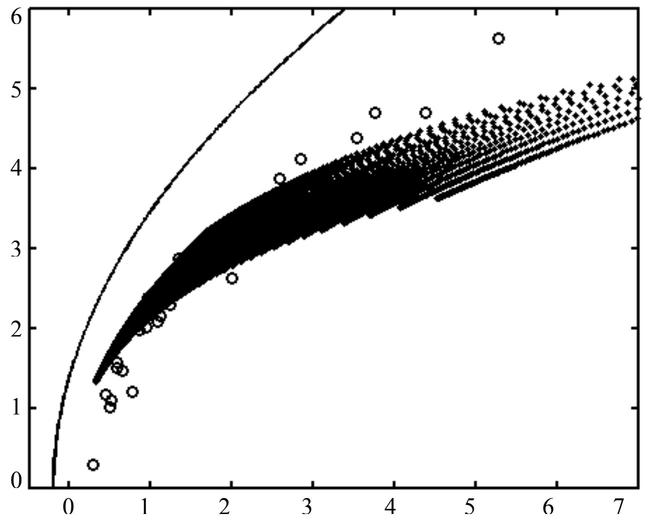

Рис. 4: Емпіричні значення коефіцієнтів ексцесу та асиметрії витрат на лікування серцево-судинних захворювань різних категорій пацієнтів у швейцарському госпіталі CHUV та їх теоретичні аналоги для параметричної сім'ї функцій щільності Фішера

Фішера та Стьюдента. Проте їх аналіз виходить за межі розгляду у даній роботі та є предметом сучасних досліджень фінансової економетрики [2-4, 8], початок яких був закладений у роботах [1, 6].

Висновки. Показано, що емпіричні значення коефіцієнтів ексцесу й асиметрії часто не відповідають можливій теоретичній області покриття даної сім'ї функцій щільності, з допомогою якої проводяться економетричні оцінки методом найбільшої правдоподібності. Така невідповідність призводить до втрати точності інтервальних оцінок параметрів розподілу, що в свою чергу веде до неточних прогнозів. Тому за додатковий критерій для оцінки адекватності моделі можна обрати здатність заданої сім'ї функцій щільності покривати емпіричну область значень коефіцієнтів асиметрії й ексцесу.

На основі запропонованої методики калібрування моделі оцінювання з використанням коефіцієнтів асиметрії й ексцесу у роботі показано, що стандартні параметричні розподіли не підходять до моделювання витрат на лікування серцево-судинних захворювань у швейцарському госпіталі CHUV і не адекватно описують прибутковість акцій на українській фондовій біржі (ПФТС).

Література

[1] *Engle R.* Auto-regressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation // Econometrica. — 1982. — Vol. 50. — P. 987-1007.

[2] *Jondeau E., Rockinger M.* Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements // Journal of Economic Dynamics & Control. — 2003. — Vol. 27. — P. 1699-1737.

[3] *Jondeau E., Rockinger M.* Gram-Charlier densities // Journal of Economic Dynamics and Control. — 2001. — Vol. 25. — P. 1457-1483.
[4] *Jondeau E., Rockinger M.* Entropy densities with an application to autoregressive conditional skewness and kurtosis // Journal of Econometrics. — 2002. — Vol. 106. — P. 119-142.
[5] *Kendall M., Stuart A.* The Advanced Theory of Statistics. Vol. I: Distribution Theory; Vol. II: Inference and Relationship. — Charles Griffin, 1977.
[6] *Mandelbrot B.* The speculation of certain speculative prices // Journal of Business. — 1963. — Vol. 35. — P. 394-419.
[7] *Manning W. G., Basu A., Mullahy J.* Generalized modeling approaches to risk adjustment of skewed outcomes data // Journal of Health Economics. — 2005. — Vol. 24. — P. 465-488.
[8] *Premarante G., Bera A. K.* Modeling asymmetry and excess kurtosis in stock return data // WP, University of Illinois. — 1999.
[9] *Widder D. V.* The Laplace Transform. — Princeton University Press, 1946.

Using Coefficients of Skewness and Kurtosis in Parametric Statistical Models

Vasyl Yeleyko, Yevhen Pentsak

The methodology of using skewness and kurtosis coefficients of random variables parametric distributions in financial and health care econometric models is considered. Fat tails and high skewness are regular characteristics of considered random variables. Visualization of normalized third and fourth moments of empirical distribution allows us to choose in an adequate way a parametric family of distributions for which respected normalized moments cover their empirical counterparts. Based on examples we show that standard parametric density functions used for excess skewness and kurtosis modelling are not able to generate empirical normalized central moments of order three and four using data set of health care expenses and assets returns on Ukrainian stock market.

Использование коэффициентов асимметрии и эксцесса в параметрических статистических моделях

Василий Елейко, Евгений Пенцак

Предлагается методика использования коэффициентов асимметрии и эксцесса параметрических распределений случайных величин в моделях финансовой екoнометрики и екoнометрики здравоохранения, для которых характерно наличие «толстых хвостов» и большого значения коэффициента асимметрии. Визуализация нормализированных моментов третьего и четвертого порядков эмпирического распределения позволяет адекватно выбрать параметрическую семью распределений, соответствующие моменты которой покрывают их эмпирические аналоги. На примерах показано, что стандартные параметрические функции плотности, которые используются для моделирования асимметрии и эксцесса, не могут генерировать эмпирических нормализированных моментов третьего и четвертого порядков для данных расходов на лечение и прибыльности акций на украинской фондовой бирже.

Отримано 27.11.06