

УДК 336.77:330.131.7

ДИСКРИМІНУЮЧА ЗДАТНІСТЬ ФУНКЦІЙ ОЦІНКИ

К. Бейнет, П. Рейчлінг, Б. Вогт

Щоб проаналізувати прогнозуючу силу оцінки, у статті обговорюються показники дискримінуючої здатності функції оцінки (або рейтингової функції). З цією ж метою, крім показників, які часто обговорюються в літературі, розглянуто ідею стохастичної домінантності. Всі перелічені критерії використано для здійснення емпіричного порівняння показників двох рейтингових агентств – Standard & Poor's та Moody's Investors Service.

Ключові слова: кредитний ризик, точність оцінювання, дискримінуюча здатність.

Вступ

Відповідно до угоди Базель II, банк може робити вибір між впровадженням підходу, що базується на внутрішній оцінці (IRB), та стандартного методу, який бере за основу результати оцінювань, проведених зовнішніми агентствами, щоб встановити мінімальну потребу в капіталі для кредитного ризику. Мета кожного підходу – оцінити можливість невиконання грошових зобов'язань позичальником. В межах рейтингової системи дебітори з подібною очікуваною вірогідністю несплати згруповані в одній рейтинговій категорії. Оцінка платоспроможності з надійним прогнозом є центральним фактором успішної діяльності банку в області кредитування.

Об'єктивність системи рейтингових оцінок стала головною темою з того часу, як органи фінансового нагляду почали приймати заяви від банків стосовно утвердження методу IRB. Кваліфікаційні перевірки вимагають визнання об'єктивності рейтингових систем перед їх схваленням. Останнім часом, з метою підтвердження функцій оцінки, у літературних джерелах обговорюється якість показників дискримінуючої здатності¹. Ці показники стосуються методів, які були розроблені близько 30 років тому і використовуються в області медицини, прогнозу погоди та в теорії виявлення сигналів. Як доповнення до існуючої літератури, яка присвячена вивченню показників дискримінуючої здатності, ми показуємо, що стохастична домінантність може також бути використана для оцінки рейтингової функції.

Статтю організовано наступним чином: у другому розділі представлено показники дискримінуючої здатності. Тут порівняння функцій оцінки за допомогою використання стохастичної домінантності буде головним питанням. У третьому розділі ми емпірично проаналізуємо рейтинги Standard & Poor's та Moody's Investors Service. В останньому розділі запропоновано висновки.

1. Показники дискримінуючої здатності

Якщо рейтингова функція демонструє високу дискримінуючу здатність, вона дає змогу розмежувати дебіторів з високим та низьким рівнем кредитоспроможності. Показники, необхідні для оцінки дискримінуючої здатності, яка обговорюється в цьому розділі, це: таблиця сполучення ознак, операційні характеристики одержувача, параметр загальної точності, стохастична тенденція та стохастична домінантність.

¹ Див. Собехарт, Кінен та Стейн (2000), Собехарт, Кінен (2001), Гайден (2002, ст. 78-95), Енгелманн, Гайден та Таше (2003), Хамерле, Раумеєр та Рош (2003).

1.1. Огляд загальних показників

Спершу розглянемо модель оцінки з двома можливими рейтингами – “невиконання зобов’язань” (дефолт) та “не невиконання зобов’язань” (не дефолт), щоб оцінити функцію оцінки. Тут ініціатор оцінювання закріплює один з двох можливих наслідків за кожним одержувачем позики в час $t = 0$, а потім спостерігає за виконанням протягом $t = 1$ (табл. 1).

Таблиця 1

Таблиця сполучення ознак

		Результати спостережень протягом $t = 1$	
		Дефолт	Не дефолт
Прогноз протягом $t = 0$	Дефолт	A	B
	Не дефолт	C	D

Для визначення дискримінуючої здатності основної функції оцінки існує серія коефіцієнтів. Коефіцієнти, які найчастіше тут використовуються, це коефіцієнт ефективності (HR) та частота появи сигналів фальшивої тривоги (FAR):

$$HR \equiv \frac{A}{A+C} \quad \text{та} \quad FAR \equiv \frac{B}{B+D}. \quad (1)$$

Таблиця сполучення придатна не лише для моделей, які просто розподіляють розряди “дефолт” чи “не дефолт”. Вона може також бути розширена для будь-якої кількості класів оцінки. Для моделей з більш ніж двома можливими класами необхідно визначити точку відсікання (мінімально прийнятний рівень доходу) (CoP), яка представляє межу, за якою позичальник розглядається як позичальник з підвищеним ризиком невиконання зобов’язань. В моделі класу багатократного оцінювання накопичений коефіцієнт ефективності HR_s для точки відсікання CoP_s зводиться до суми одиничних коефіцієнтів ефективності найгірших класів оцінки $1, \dots, s$. Накопичений коефіцієнт появи сигналів фальшивої тривоги FAR_s обчислюється аналогічно.

Таблиця сполучення лише дає змогу оцінити можливість класифікації функції оцінки відносно обраної точки відсікання. Загальна дискримінуюча здатність функції оцінки між дебіторами з низьким та високим рівнями кредитоспроможності, яка згрупована за всіма точками відсікання, представлена за допомогою операційних характеристик позичальника.

1.1.1. Операційні характеристики позичальника та площина під відрізком кривої

Операційні характеристики позичальника (ROC) було розроблено у 1950-х роках, щоб перевірити міцність радіосигналів, на які вплинув шум. Криву ROC одержуємо за допомогою нанесення коефіцієнтів ефективності проти коефіцієнтів появи сигналів фальшивої тривоги для всіх можливих точок відсікання. На рисунку 1 показано криву ROC¹.

Площина під кривою ROC (жирна лінія на рис. 1) має назву “площина під відрізком кривої” (AUC). Вона оцінює дискримінуючу здатність через порівняння з площиною нижче функції точного оцінювання (тонка лінія). Значення 50% AUC знаходиться у відповідності з випадковим експериментом з метою прогнозування неплатоспроможності (пунктирна лінія).

¹ Див. Собехарт, Кінен (2001) та Енгелманн, Гайден та Таче (2003).

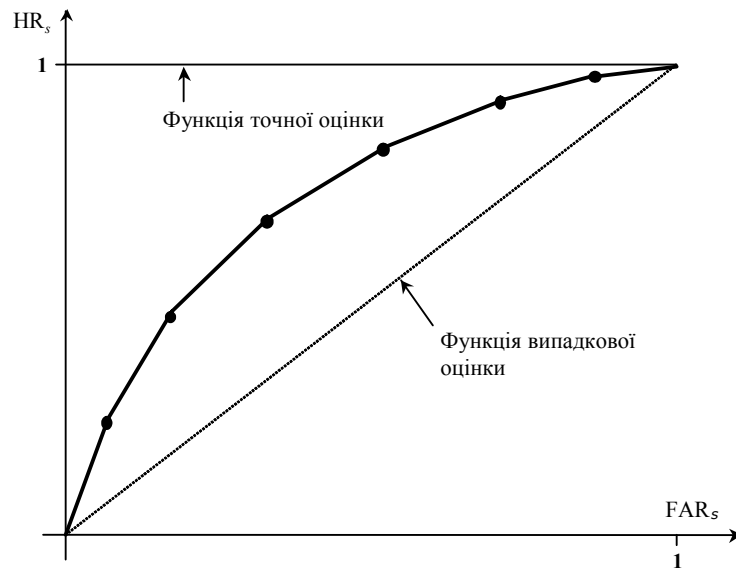


Рис. 1. Операційні характеристики позичальника

Здатність функції оцінки приписати найнижчу можливу категорію дебіторам, які стали неплатоспроможними, та найкращу – тим, що зберегли свою кредитоспроможність, не оцінюється за допомогою AUC. Важливою ознакою моделі ROC є те, що менш кредитоспроможні дебітори отримують гірші рейтинги, ніж дебітори з високим рівнем платоспроможності.

1.1.2. Параметр загальної точності та коефіцієнт точності

Параметр загальної точності (CAP) виникає згідно з кривою Лоренца. Він оцінює здатність рейтингової функції помістити дебіторів з низьким рівнем кредитоспроможності до найнижчого класу оцінки.

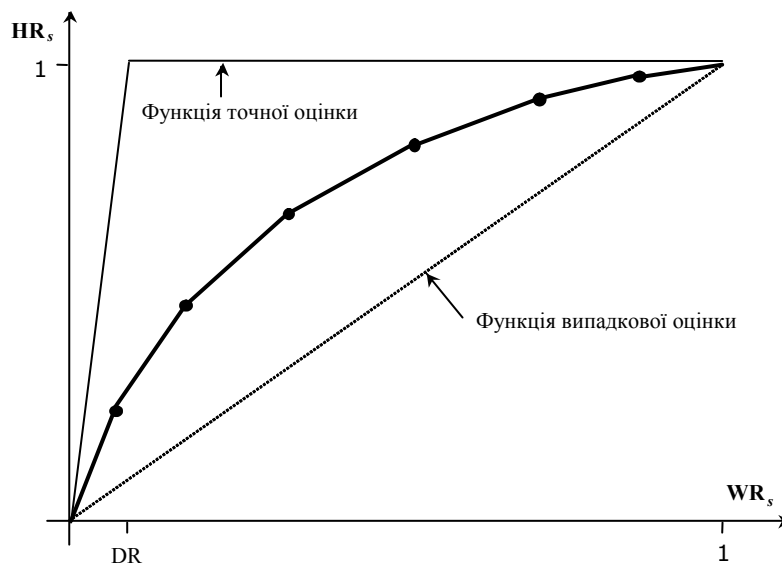


Рис. 2. Параметр загальної точності

Криву CAP отримуємо за допомогою сортування всіх дебіторів згідно з їхнім рейтингом. Далі обчислюємо, яка категорія дебіторів з найгіршими рейтингами WR_s демонструє коефіцієнт ефективності неплатоспроможності. Ідеальна крива CAP (тонка крива на рисунку 2) визначається штрафною процентною ставкою (DR) портфеля.

Разом з параметром загальної точності показник точності AR, який обчислюється у відповідності зі стандартизованим коефіцієнтом Джині кривої Лоренца, служить для оцінки рейтингової функції. Цей коефіцієнт визначає відношення між двома площинами: перша – площа між лінією експериментальної функції оцінки (жирна лінія на рисунку 2) та лінією функції випадкового оцінювання (пунктирна лінія), друга – площа між функцією точної оцінки (тонка лінія) та функцією випадкового оцінювання¹.

На відміну від коефіцієнта Джині показник точності може набути негативного значення. Коефіцієнт точності та площа під відрізком кривої можуть переходити один в одного, як показано нижче²:

$$AR = 2 \cdot AUC - 1. \quad (2)$$

1.1.3. Стохастична тенденція

Стохастична тенденція також може бути використана для оцінки того, чи відрізняються один від одного розподіли (класифікації) коефіцієнта ефективності та коефіцієнта появи сигналів фальшивої тривоги. В нашій дискретній моделі розподіл коефіцієнта FAR наближається до більшого стохастичного значення, ніж класифікація HR, якщо застосовується наступна нерівність:

$$RE \equiv \sum_s \frac{HR_s + HR_{s-1}}{2} \cdot (FAR_s - FAR_{s-1}) > \frac{1}{2} \quad \text{де } HR_0 = FAR_0 \equiv 0. \quad (3)$$

З відносним ефектом (RE) вище 0.5 компанії, які залишились платоспроможними, були поміщені до категорій високої оцінки, а неплатоспроможні дебітори – до класів низької оцінки. RE та AUC в сумі дають результат FAR на клас оцінки та накопичений HR. Тому обидва поняття кількісно оцінюють площину під кривою ROC.

Нехай S^d та S^{nd} означають рейтинг дебіторів, які стали неплатоспроможними та залишились платоспроможними відповідно. Тоді, у імовірнісній інтерпретації, коефіцієнти ефективності та коефіцієнти появи сигналів фальшивої тривоги є наступні:

$$HR_s = Prob(S^d \leq s) \quad \text{ма} \quad FAR_s = Prob(S^{nd} \leq s). \quad (4)$$

У випадку незалежного оцінювання це формулювання дає наступну інтерпретацію відносного ефекту та площини під відрізком кривої відповідно³:

$$RE = AUC = Prob(S^d < S^{nd}) + \frac{1}{2} \cdot Prob(S^d = S^{nd}). \quad (5)$$

Площина під відрізком кривої відповідає імовірності, що неплатоспроможна компанія отримує гіршу оцінку, ніж дебітор, який залишився кредитоспроможним. Значимо, що відносний ефект не реагує на трансформацію, що зберігає порядок⁴. Тому відносний ефект, площа під кривою та показник точності притаманні показникам порядкового оцінювання. Незалежно від того, чи функція оцінки забезпечує визначення результатів чи імовірність

¹ Див. Собехарт, Кінен та Стейн (2000), Енгелманн, Гайден та Таче (2003).

² Див. Енгелманн, Гайден та Таче (2003). Зауважимо, що ці показники з фіксованими рейтингами, залежать від штрафної процентної ставки кредитного портфеля. Див. Гамерле, Раумеер та Рош (2003) та Собехарт та Кінен (2004).

³ Див. Бамбер (1975).

⁴ Див. Бамбер (1975).

невиконання грошових зобов'язань, ці показники приходять до достовірного результату, поки теоретична імовірність дефолту зростає з низьким рейтинговим показником.

В якості проміжного результату маємо наступне: площа під відрізком кривої ROC дорівнює відносному ефекту стохастичної тенденції. Обидва поняття перебувають в лінійній залежності з показником точності. Тому показники дискримінуючої здатності, яка обговорюється, поки що дають ідентичні результати.

1.2. Стохастична домінантність

Стохастична домінантність першого порядку оцінює сумарну вірогідність двох випадкових змінних. У нашій моделі стохастична домінантність може бути застосована по відношенню до розподілів HR та FAR однієї функції оцінки або розподілів HR двох функцій оцінки.

У випадку однієї рейтингової функції маємо стохастичну домінантність першого порядку класифікації FAR над класифікацією HR, якщо використовується наступна нерівність:

$$FAR_s \leq HR_s \quad \forall_s \quad (6)$$

і якщо ця нерівність точно виконується принаймні для одного класу оцінки s . Тоді крива ROC, яка є предметом спостереження, знаходиться над діагоналлю, що на рисунку 1, результатом чого є площа під відрізком кривої понад 50%.

Для порівняння двох рейтингових функцій розглянемо коефіцієнт ефективності $HR(s) = HR_s$ та коефіцієнт появи сигналів фальшивої тривоги $FAR(s) = FAR_s$. Далі крива ROC представлена вектором $(HR(s), FAR(s))$. В контексті стохастичної домінантності першого порядку функція оцінки R займає домінуюче положення по відношенню до рейтингової функції T у випадку ідентичних FAR, якщо наступна нерівність зберігається та точно виконується принаймні для одного рейтингового класу s :

$$HR_T(s) \leq HR_R(s) \quad \forall_s \quad (7)$$

Оскільки FAR – це монотонне перетворення показників оцінки, нерівність (7) зберігається для всіх коефіцієнтів FAR. Тому у випадку ідентичних FAR крива ROC функції оцінки R проводиться над кривою функції оцінки T . В результаті, значення AUC рейтингової функції R виявляється вищим, ніж значення функції T .

За відсутності стохастичної домінантності першого порядку доцільним буде спробувати застосувати показник стохастичної домінантності другого порядку. Стохастична домінантність другого порядку класифікації FAR займає домінуюче положення по відношенню до розподілу HR, якщо наступна нерівність зберігається та чітко виконується принаймні для одного класу оцінки i :

$$\sum_{s=1}^i FAR(s) \leq \sum_{s=1}^i HR(s) \quad \forall i \quad (8)$$

З цієї домінантності не обов'язково слідує, що площа під відрізком кривої лежить над 50%, як доведено прикладом з таблиці 2. Приклад побудовано у такий спосіб, що хоча й має місце стохастична домінантність другого порядку класифікації FAR над розподілом HR, значення AUC 47% обчислюється.

Стохастична домінантність другого порядку для коефіцієнтів ефективності

Кількісний показник	Емітенти	Невиконання зобов'язань	HR _s	FAR _s	$\sum_{s=1}^i \text{HR}_s$	$\sum_{s=1}^i \text{FAR}_s$
1	160	100	33 %	20 %	0.33	0.20
2	40	30	43 %	23 %	0.77	0.43
3	200	30	53 %	80 %	1.30	1.23
4	200	140	100 %	100 %	2.30	2.23

З цього прикладу ми бачимо, що загальні показники дискримінуючої здатності зосереджені на коефіцієнтах FAR та HR. З економічної точки зору це не повинно привести до того ж результату, що й аналіз імовірності невиконання зобов'язань у кожному класі оцінки. Наприклад, якщо не брати до уваги різні суми кредиту, альтернативні витрати та кредитні спреди, то використовуючи функцію оцінки з таблиці 2а, стратегія розподілу кредитних ресурсів може допомогти уникнути витрат, пов'язаних з невиконанням грошових зобов'язань 100 та 130 кредиторами, якщо заявки про надання позики від клієнтів з показниками 1 та 2, відповідно, не приймаються. Застосовуючи функцію випадкового оцінювання з однаковими коефіцієнтами FAR, можна побачити, що лише 60 чи скоріше 70 заявок клієнтів, які стали банкрутами, відхиляються.

В контексті стохастичної домінантності другого порядку ми знову порівнюємо рейтингові функції R та T з однаковими FAR. Функція оцінки R займає домінуюче положення по відношенню до рейтингової функції T , якщо наступна нерівність зберігається та чітко виконується принаймні для одного класу оцінки i :

$$\sum_{s=1}^i \text{HR}_T(s) \leq \sum_{s=1}^i \text{HR}_R(s) \quad \forall i. \quad (9)$$

Якщо ми порівняємо функцію оцінки з таблиці 2 та функцію випадкового оцінювання з тими самими коефіцієнтами FAR, то функція випадкової оцінки продемонструє вище значення AUC 0.5. Проте в плані стохастичної домінантності другого порядку перевага віддається функції оцінки з таблиці 2. Наші приклади з ідентичними коефіцієнтами FAR показують, що поняття площини під відрізком кривої та показника точності, відповідно, заперечують показник стохастичної домінантності другого порядку.

2. Емпіричні результати

Метою нашого емпіричного дослідження є розкриття двох питань. По-перше, яку дискримінуючу здатність демонструють оцінки агентств Standard & Poor's (S&P) та Moody's Investors Service (Moody's)? Оскільки ці компанії є найбільш відомими у всьому світі рейтинговими агентствами, ми очікуємо, що вони володіють високими значеннями AUC та AR. По-друге, чи займає функція оцінки одного агентства домінуюче положення по відношенню до функції оцінки іншої компанії? Щоб відповісти на ці питання, аналізуємо період з 1982 по 2001 рік. У своєму розпорядженні ми мали кількість емітентів, а також штрафних процентних ставок за цей строк та сім категорій шкали оцінювання S&P (AAA, AA ... CCC) та Moody's (Aaa, Aa ... C). Базу даних склали звіти цих агентств¹.

¹ Див. Moody's Investors Service (2002) та Standard & Poor's (2002). Кількість емітентів Moody's була забезпечена Moody's KMV.

Дані Moody's включають інформацію про емітентів довгострокових боргових зобов'язань з секторів промисловості, транспорту, комунальних послуг та фінансових установ. Дані про емітентів структурованих фінансових продуктів та державних емітентів відсутні. Інформація S&P також надійшла від емітентів довгострокових боргових зобов'язань, головним чином з секторів промисловості, комунальних послуг страхування та фінансових установ. І знову дані про емітентів структурованих фінансових продуктів, державних банків-емітентів та емітентів, чий показник базувались лише на державній інформації, були виключені. На рисунку 3 продемонстровано значення AUC для нашої вибірки. Всі обчислені значення AUC суттєво відрізняються від 0.5 за статистикою Манна-Вітні (всі значення p нижчі за 0.1%).

Функції оцінки S&P та Moody's демонструють дискримінуючу здатність для нашої вибірки. Середнє значення AUC вибірки агентства Moody's – 90.9%, а середнє значення для S&P складає 90.5%. Згідно з результатами перевірки знаків, значення AUC для Moody's є значно вищими, ніж для S&P (значення p – нижче 6%). Всі встановлені криві ROC розташовані над діагоналлю, що на рисунку 1. У кожному розглянутому тут випадку ми виявили стохастичну домінуючість першого порядку класифікації FAR над розподілом HR.

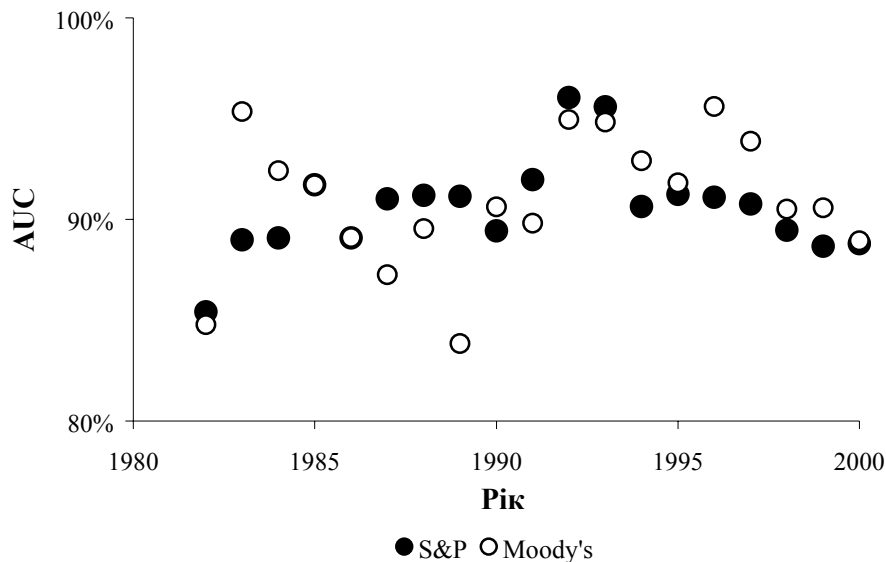


Рис. 3. Значення AUC для рейтингів S&P та Moody's (1982-2001)

Дещо інший результат виникає, коли порівнюються лише класифікації HR. В 17 (з двадцяти) роках ранжування рейтингових функцій є можливим згідно зі стохастичною домінуючістю другого порядку без суттєвої переваги котрогось з рейтингових агентств. В 1985, 1986 та 1990 роках ми спостерігали різну градацію стохастичної домінуючість HR та показника AUC. Оскільки відмінності між коефіцієнтами FAR здаються незначними, це підтверджує наш теоретичний доказ з розділу 2.2. В таблиці 3 запропоновано перегляд.

Стохастична домінантність другого порядку та домінантність AUC

Рік	Стохастична домінантність другого порядку HR	Домінантність AUC	Різниця FAR ¹
1982	–	S&P	4.1 %
1983	Moody's	Moody's	1.5 %
1984	–	Moody's	2.7 %
1985	S&P	Moody's	2.1 %
1986	S&P	Moody's	2.9 %
1987	S&P	S&P	3.6 %
1988	S&P	S&P	3.3 %
1989	S&P	S&P	2.8 %
1990	S&P	Moody's	2.2 %
1991	S&P	S&P	1.8 %
1992	S&P	S&P	2.4 %
1993	S&P	S&P	2.4 %
1994	–	Moody's	2.5 %
1995	Moody's	Moody's	2.6 %
1996	Moody's	Moody's	2.4 %
1997	Moody's	Moody's	2.6 %
1998	Moody's	Moody's	2.1 %
1999	Moody's	Moody's	1.7 %
2000	Moody's	Moody's	2.7 %
2001	Moody's	Moody's	3.4 %

Висновок

Крім загальних показників, необхідних для оцінки дискримінуючої здатності функцій оцінки, як то площа під відрізком кривої, показник точності та відносний ефект, ми запропонували стохастичну домінантність. Хоча стохастична домінантність першого порядку веде до тих самих наслідків, що й загальні показники, стохастична домінантність другого порядку, виявляється, дає інші результати.

Наш емпіричний аналіз забезпечує наступний результат. Функції оцінки Standard & Poor's та Moody's Investors Service володіють дискримінуючою здатністю, як і очікувалось, для 1982 та 2001 років. На жаль, ми не маємо змоги дослідити постійну домінантність певного рейтингового агентства.

Список використаних джерел

1. Bamber, D. The Area above the Ordinal Dominance Graph and the Area below the Receiver Operating Characteristic Graph // Journal of Mathematical Psychology, 1975. – №12. – pp. 387-415.
2. Engelmann, B., E. Hayden, D. Tasche. Testing Rating Accuracy // Risk, 2003. – №16. – pp. 82-86.

¹ Різниця FAR $\equiv \sqrt{\frac{1}{6} \sum_{s=1}^6 (FAR_s^{S\&P} - FAR_s^{Moody's})^2}$

3. Hamerle, A., R. Rauhmeier, D. Roesch. Uses and Misuses of Measures for Credit Rating Accuracy // Working Paper, University of Regensburg, 2003.
4. Hayden, E. Modeling an Accounting-Based Rating System for Austrian Firms // Dissertation Thesis, University of Vienna, 2002.
5. Moody's Investors Service. Default & Recovery Rates of Corporate Bond Issuers // Special Comment, 2002 – February.
6. Sobehart, J., S. Keenan. Measuring Default Accurately // Risk, 2001. – №14.– pp. 31-33.
7. Sobehart, J., S. Keenan. The Score for Credit // Risk, 2004. – №17.– pp. 54-58.
8. Sobehart, J., S. Keenan, R. Stein. Validation Methodologies for Default Risk Models // Credit, 2000. – №5. – pp. 51-56.
9. Standard & Poor's. Ratings Performance 2001 // Special Report, 2002 – February.

Отримано 01.02.2007