

*И.М. Никольский<sup>1</sup>, К.К. Фурманов<sup>2</sup>*

## **АСИМПТОТИЧЕСКИЕ КОЭФФИЦИЕНТЫ РАНГОВОЙ КОРРЕЛЯЦИИ МЕЖДУ ИСТИННЫМИ ПОКАЗАТЕЛЯМИ НЕЭФФЕКТИВНОСТИ И ИХ ОЦЕНКАМИ В ПРОСТОЙ МОДЕЛИ СТОХАСТИЧЕСКОЙ ГРАНИЦЫ**

### **Введение**

Достижение эффективности использования ресурсов — одна из основных целей управления хозяйствующими субъектами. Чтобы выяснить, в какой мере эта цель достигается, нужно измерить эффективность субъектов, принимающих решения о распределении ресурсов. В настоящее время для измерения эффективности широко применяются методы анализа данных – в основном, анализ оболочки данных (Data Envelopment Analysis, DEA) и методы регрессионного анализа. Примером применения обоих подходов может послужить статья[1].

Среди регрессионных моделей большое распространение получили модели стохастической границы производственных возможностей. Впервые такая модель была предложена в статьях [2] и [3], после чего появилось множество её разновидностей, обзор которых содержится в работе [4]. Анализ стохастической границы широко применяется в различных сферах: для моделирования эффективности промышленных предприятий [5], [6], кредитных организаций [7], [8], [9], учебных заведений [10], [1] и целых регионов [11], [12], [13].

В настоящей статье выводятся формулы для коэффициентов ранговой корреляции Кендалла и Харрелла между истинными показателями эффективности субъектов принятия решения и их оценками в модели стохастической границы с показательным распределением показателя неэффективности, предложенной в [2] и [3]. Формулы позволяют рассчитать асимптотические (при увеличении объёма выборки) значения коэффициентов ранговой корреляции, исходя из оценок модели, и таким образом оценить ранжирующую точность

---

<sup>1</sup> Москва, ф-т ВМК МГУ имени М.В. Ломоносова, доцент, e-mail: inikolsky@cs.msu.ru

<sup>2</sup> Москва, ЦЭМИ РАН, лаборатория прикладной эконометрики, старший научный сотрудник, e-mail: kfurmanov@hse.ru.

модели, её способность выделять относительно эффективные и неэффективные субъекты.

### Обзор литературы и описание базовой модели стохастической границы

Модель стохастической границы была предложена в работах двух исследовательских коллективов [3], [2] как регрессионная модель, в которой объясняемая величина отражала объём производства предприятия, а объясняющие — затраты факторов производства. Особенность модели состояла в разделении ошибки регрессии на две компоненты: нормально распределённую величину, отражающую «стохастический шок» (индивидуальные различия в границе производственных возможностей между субъектами), и неотрицательную величину, отражающую неэффективность субъекта (отклонение наблюдаемого выпуска от границы производственных возможностей). Модель представлялась следующим уравнением регрессии:

$$y_i = f(x_i, \beta) + v_i - u_i, \quad (1)$$

где  $y_i$  — объём выпуска предприятия  $i$  (обычно в логарифмированном виде),  $x_i$  — вектор затрат факторов производства,  $\beta$  — вектор параметров производственной функции,  $f(x_i, \beta)$  — функция, отражающая максимально достижимый выпуск предприятия при заданных затратах  $x_i$  без учёта стохастического шока,  $v_i$  — стохастический шок,  $u_i$  — величина, отражающая неэффективность предприятия  $i$ . В качестве производственной функции  $f(x_i, \beta)$  на практике часто используется функция Кобба-Дугласа или транслог-функция [14].

Из участвующих в уравнении (1) величин две ( $x_i$  и  $\beta$ ) детерминированы, остальные — случайны. В исходной модели делались следующие предпосылки о компонентах случайной ошибки: шоки  $v_i$  нормально распределены с нулевым математическим ожиданием и одинаковой для всех предприятий дисперсией, компоненты неэффективности также одинаково распределены для всех предприятий, но их распределение либо экспоненциальное, либо полунормальное. В дальнейшем возникло множество модификаций модели с ослаблением изначальных предпосылки — обзор содержится в статье [4].

Модель (1) нашла применение не только при анализе производственных предприятий, так что в дальнейшем наряду со словом «предприятие» будем использовать выражение «субъект принятия решения».

В исходных статьях [3], [2] речь шла об оценивании параметров распределения компоненты неэффективности, из которых можно, в частности, получить оценку среднего уровня неэффективности в отрасли.

Оценка для неэффективности отдельных субъектов были впервые предложена в работе [15], но распространение получила более поздняя её версия, представленная в статье [16] и называемая JLMS-оценкой по фамилиям авторов. Другой вариант, называемой ВС-оценкой, изложен в статье [17]. Отмечается, что в целом два метода оценивания — JLMS и ВС — дают практически идентичные результаты [4].

Исследование свойств ранжировок субъектов по эффективности началось недавно. Основополагающей можно назвать статью [18], где исследовано распределение ранга отдельного субъекта. В статьях [14] и [19] рассматриваются ранжировки целиком, точнее согласованность ранжировок субъектов по эффективности, полученных в результате оценивания модели стохастической границы, и «истинных» ранжировок (по ненаблюдаемым на практике значениям  $u_i$ ). В работе [14] представлены результаты серии статистических экспериментов, показывающих, что коэффициенты ранговой корреляции Спирмена и Харрелла между истинными компонентами неэффективности  $u_i$  и их JLMS-оценками зависят почти исключительно от соотношения дисперсий стохастического шока и неэффективности. Аналитическое обоснование этому дано в статье [19], где показано, что коэффициент Харрелла между истинными компонентами неэффективностями и их оценками асимптотически (при увеличении объёма выборки до бесконечности) зависит только от распределений компонент ошибки регрессии. Применительно к исходной модели [3] с нормально распределёнными шоками и либо экспоненциальным, либо полунормальным распределением неэффективностей это означает, что коэффициент Харрелла зависит только от дисперсий этих компонент, так как дисперсии полностью задают их распределения. Для случая полунормальной неэффективности в статье [19] выведено приближённое выражение асимптотического коэффициента Харрелла через дисперсии шоков и неэффективностей. Таким образом, получена мера ранжирующей способности модели стохастической границы, позволяющая по состоятельно оцениваемым параметрам распределений выяснить, насколько модель позволяет отличать относительно эффективные предприятия от относительно неэффективных.

В настоящей статье приводится вывод выражений для коэффициента Харрелла и коэффициента ранговой корреляции Кендалла для модели стохастической границы с нормально распределённым шоком и экспоненциально распределённой неэффективностью.

## Измерение согласованности ранжировок

Для эмпирической литературы характерно сравнение оценок эффективности субъектов, полученных разными моделями и методами, для проверки устойчивости выводов к выбору инструментария. По-видимому, самая распространённая мера согласованности — коэффициент ранговой корреляции Спирмена (см., например, [20], [4]). К сожалению, этот подход можно применить для измерения ранжирующей способности только на искусственно сгенерированных данных, на практике же рассчитать коэффициент Спирмена невозможно, так как истинные показатели неэффективности неизвестны — иначе не стояла бы сама задача оценивания. В работе [10] показано, что возможно косвенное, аналитическое изучение точности ранжировок. При этом коэффициент ранговой корреляции выводится как функция от параметров модели стохастической границы. Для этой цели, однако, коэффициент Спирмена неудобен.

В статье [21] рассматриваются ранговые коэффициенты корреляции, удобные как для аналитического исследования, так и для практического использования при анализе реальных данных, так как они связаны с ясными и интерпретируемыми параметрами совместного распределения случайных величин, связь между которыми подлежит изучению. Наиболее известный из них — коэффициент ранговой корреляции Кендалла, определяемый следующим образом. Пусть пары случайных величин  $(X_1, Y_1)$  и  $(X_2, Y_2)$  имеют одинаковые совместные распределения и независимы. Подразумевается независимость именно между парами, величины  $X_i$  и  $Y_i$  могут быть зависимы. Коэффициент ранговой корреляции Кендалла  $\tau_{XY}$  определяется равенством:

$$\tau_{XY} = E(\text{sign}(X_1 - X_2)\text{sign}(Y_1 - Y_2)),$$

где  $E(\dots)$  — математическое ожидание, а  $\text{sign}$  — функция знака.

Коэффициент ранговой корреляции Сомерса  $D_{YX}$  — некоммутативная мера связи, определяемая формулой

$$D_{YX} = P(Y_1 < Y_2 | X_1 < X_2) - P(Y_1 > Y_2 | X_1 < X_2)$$

и связанная с коэффициентом Кендалла равенством  $D_{YX} = \frac{\tau_{XY}}{\tau_{XX}}$  (см. [22], [21]).

Обратим внимание на то, что для непрерывной случайной величины  $X$  выполняется равенство  $\tau_{XX} = 1$ , поэтому  $D_{YX} = \tau_{XY}$ . Если же обе величины  $X$  и  $Y$  непрерывны, то коэффициент Сомерса коммутативен:  $D_{YX} = D_{XY}$ .

Тесно связан с коэффициентом Сомерса коэффициент Харрелла  $C_{YX}$  (он же индекс конкордации Харрелла,  $C$ -индекс и просто  $C$  Харрелла, см. [23]), определяемый равенством<sup>3</sup>

$$C_{YX} = P(Y_1 < Y_2 | X_1 < X_2).$$

В отличие от коэффициентов Кендалла и Сомерса, принимающих значения из интервала  $[-1; 1]$ , индекс конкордации лежит в пределах  $[0; 1]$ , где 1 — полная согласованность ранжировок, а 0 — полное расхождение ранжировок. Отсутствию связи между рангами соответствует значение 0.5.

В дальнейшем нам будет интересен только случай непрерывных случайных величин  $X$  и  $Y$ . В этом случае коэффициенты Харрелла и Сомерса коммутативны и связаны с коэффициентом Кендалла равенством:

$$\tau_{XY} = D_{XY} = D_{YX} = 2C_{YX} - 1 = 2C_{XY} - 1. \quad (2)$$

Таким образом, в непрерывном случае достаточно найти значение одного из этих коэффициентов, чтобы рассчитать остальные.

#### Асимптотические коэффициенты ранговой корреляции в модели стохастической границы

Пусть имеется выборка  $(x_1, y_1), \dots, (x_n, y_n)$ , где  $x_i$  — детерминированный вектор объясняющих переменных (затрат факторов производства, если речь идёт о моделировании эффективности производства), а  $y_i$  — случайная величина, объясняемая переменная (объём производства) для субъекта  $i$ . Значения объясняемой и объясняющих переменных связаны уравнением регрессии (1). Стохастические шоки  $v_1, \dots, v_n$  одинаково распределены, компоненты, отражающие неэффективность субъектов,  $u_1, \dots, u_n$  также одинаково распределены. Величины  $v_1, \dots, v_n, u_1, \dots, u_n$  независимы. Производственная функция  $f(x, \beta)$  непрерывна по параметрам  $\beta$ .

Пусть имеется состоятельная оценка  $\hat{\beta}$  для вектора параметров производственной функции (это может быть, например, оценка метода максимального правдоподобия) и соответствующие JLMS-оценки неэффективностей  $\hat{u}_1, \dots, \hat{u}_n$ . Показано, что асимптотический коэффициент Харрелла согласованности ранжировок истинных неэффективностей  $u_1, \dots, u_n$  и их JLMS-оценок выражается следующим образом (см. [19], с. 132–133):

$$C_{\hat{u}, u}^{asy} = \lim_{n \rightarrow \infty} P(\hat{u}_1 < \hat{u}_2 | u_1 < u_2) = P(v_1 - u_1 > v_2 - u_2 | u_1 < u_2). \quad (3)$$

<sup>3</sup> Можно встретить различные определения коэффициента Харрелла, но эти различия не имеют значения для непрерывных случайных величин и не важны для настоящей статьи.

Индексы 1 и 2 здесь поставлены произвольно — в качестве наблюдений 1 и 2 могут фигурировать любые наблюдения из выборки, так как соответствующие случайные величины все независимы и одинаково распределены.

Из выражения (3) следует, что асимптотический коэффициент Харрелла определяется только распределением стохастических шоков и неэффективностей. Предположим теперь, что шоки распределены нормально,  $v_i \sim N(0, \sigma_v^2)$ , а неэффективности имеют экспоненциальное распределение с дисперсией  $\sigma_u^2$ , так что их функции плотности имеют вид  $f_u(t) = \sigma_u^{-1} \exp(-t\sigma_u^{-1})$ ,  $t \geq 0$ . Эта предпосылка используется в статье [2], а в работе [3] допускаются два варианта семейств распределения неэффективности: экспоненциальное и полунормальное.

Введём случайные величины  $\eta = v_1 - v_2$  и  $\xi = u_1 - u_2$  и выразим через них асимптотический коэффициент Харрелла:

$$C_{u,u}^{asy} = P(\eta - \xi > 0 | \xi < 0). \quad (4)$$

Функцию плотности случайной величины  $\xi$  можно найти по формуле свёртки:

$$f_\xi(x) = \int_{-\infty}^{+\infty} f_u(t)f_u(x+t)dt = \int_0^{+\infty} \sigma_u^{-1} \exp(-t\sigma_u^{-1})f_u(x+t)dt. \quad (5)$$

В случае  $x > 0$   $f_u(x+t) = \sigma_u^{-1} \exp(-(x+t)\sigma_u^{-1})$  и выражение (5) приобретает следующий вид:

$$f_\xi(x) = \int_0^{+\infty} (\sigma_u^{-1} \exp(-t\sigma_u^{-1}))(\sigma_u^{-1} \exp(-(x+t)\sigma_u^{-1}))dt = \int_0^{+\infty} \sigma_u^{-2} \exp(-(x+2t)\sigma_u^{-1})dt.$$

После интегрирования получим

$$f_\xi(x) = \frac{-\sigma_u^{-1}}{2} \exp(-x\sigma_u^{-1}) \exp(-2t\sigma_u^{-1})|_0^{+\infty} = \frac{\sigma_u^{-1}}{2} \exp(-x\sigma_u^{-1}).$$

Для случая  $x < 0$  аналогично выводится выражение

$$f_\xi(x) = \frac{\sigma_u^{-1}}{2} \exp(x\sigma_u^{-1}).$$

Объединяя два случая, получаем, что величина  $\xi$  имеет распределение Лапласа с функцией плотности

$$f_\xi(x) = \frac{\sigma_u^{-1}}{2} \exp(-|x|\sigma_u^{-1}).$$

Вернёмся к выражению (4). Отметим, что фигурирующие в нём величины  $\eta$  и  $\xi$  независимы, поэтому условное распределение  $\eta$  при условии  $\{\xi < 0\}$  совпадает с безусловным. Учитывая, что  $\eta = v_1 - v_2$ , а величины  $v_1$  и  $v_2$  нормально распределены и независимы, получаем  $\eta \sim N(0, 2\sigma_v^2)$ . Найдём условную плотность величины  $\xi$ :

$$f_{\xi}(x|\xi < 0) = \frac{f_{\xi}(x)}{P(\xi < 0)} = \sigma_u^{-1} \exp(x\sigma_u^{-1}), x < 0.$$

Таким образом, при условии  $\xi < 0$  величина  $-\xi$  экспоненциально распределена, а условное распределение разности  $\eta - \xi$  есть распределение суммы нормальной и экспоненциальной случайных величин, известное также как экспоненциально модифицированное распределение Гаусса [24]. Функция распределения такой величины имеет вид

$$F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right) - \exp\left(\frac{\lambda}{2}(2\mu + \lambda\sigma^2 - 2x)\right) \Phi\left(\frac{x - \mu - \lambda\sigma^2}{\sigma}\right),$$

где  $\Phi$  — функция стандартного нормального распределения,  $\mu$  и  $\sigma^2$  — параметры нормального слагаемого,  $\lambda$  — параметр экспоненциального слагаемого. В нашем случае  $\mu = 0, \sigma^2 = 2\sigma_v^2, \lambda = \sigma_u^{-1}$ , так что функция условного распределения величины  $\eta - \xi$  приобретает следующий вид:

$$F_{\eta-\xi}(x|\xi < 0) = \Phi\left(\frac{x}{\sigma_v\sqrt{2}}\right) - \exp\left(\frac{\sigma_v^2}{\sigma_u^2} - \frac{x}{\sigma_u}\right) \Phi\left(\frac{x - 2\sigma_v^2\sigma_u^{-1}}{\sigma_v\sqrt{2}}\right). \quad (6)$$

Сопоставляя формулы (4) и (6), получаем выражение для асимптотического коэффициента Харрелла:

$$C_{u,u}^{asy} = P(\eta - \xi > 0|\xi < 0) = 1 - F_{\eta-\xi}(0|\xi < 0) = 1 - \Phi(0) + \exp\left(\frac{\sigma_v^2}{\sigma_u^2}\right) \Phi\left(\frac{-2\sigma_v^2\sigma_u^{-1}}{\sigma_v\sqrt{2}}\right),$$

т.е.

$$C_{u,u}^{asy} = \frac{1}{2} + \exp\left(\frac{\sigma_v^2}{\sigma_u^2}\right) \Phi\left(-\frac{\sigma_v\sqrt{2}}{\sigma_u}\right).$$

Таким образом, согласованность ранжировок определяется отношением дисперсий стохастического шока и неэффективности. Введём обозначение  $\rho = \frac{\sigma_v^2}{\sigma_u^2}$  и перепишем асимптотический коэффициент Харрелла как функцию от  $\rho$ :

$$C_{u,u}^{asy} = \frac{1}{2} + \exp(\rho) \Phi(-\sqrt{2\rho}). \quad (7)$$

Альтернативное выражение можно получить через величину  $\theta = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2} = \frac{1}{1 + \rho}$  — долю дисперсии неэффективности в дисперсии суммарной случайно ошибки:

$$C_{u,u}^{asy} = \frac{1}{2} + \exp\left(\frac{1-\theta}{\theta}\right) \Phi\left(-\sqrt{\frac{2-2\theta}{\theta}}\right). \quad (8)$$

Такое выражение удобнее для графического представления, так как  $\theta$  принадлежит ограниченному множеству  $[0; 1]$ , и на графике можно

отобразить значения коэффициента при всех возможных соотношениях дисперсий  $\sigma_v^2$  и  $\sigma_u^2$ .

Из формул (2), (7) и (8) следуют выражения для асимптотического коэффициента корреляции Кендалла через отношение дисперсий шока и неэффективности  $\rho$  и через долю дисперсии неэффективности в дисперсии суммарной ошибки  $\theta$ :

$$\tau_{\hat{u},u}^{asy} = 2C_{\hat{u},u}^{asy} - 1 = 2\exp(\rho)\Phi(-\sqrt{2\rho}) \quad (9)$$

$$\tau_{\hat{u},u}^{asy} = 2\exp\left(\frac{1-\theta}{\theta}\right)\Phi\left(-\sqrt{\frac{2-2\theta}{\theta}}\right) \quad (10)$$

Как уже упоминалось в предыдущем разделе, коэффициент Сомерса в данном случае равен коэффициенту Кендалла. На рис. 1 представлены графики зависимости коэффициентов Харрелла и Кендалла от параметра  $\theta$ . По графику видно, что согласованность ранжировок растёт при увеличении доли дисперсии неэффективности.

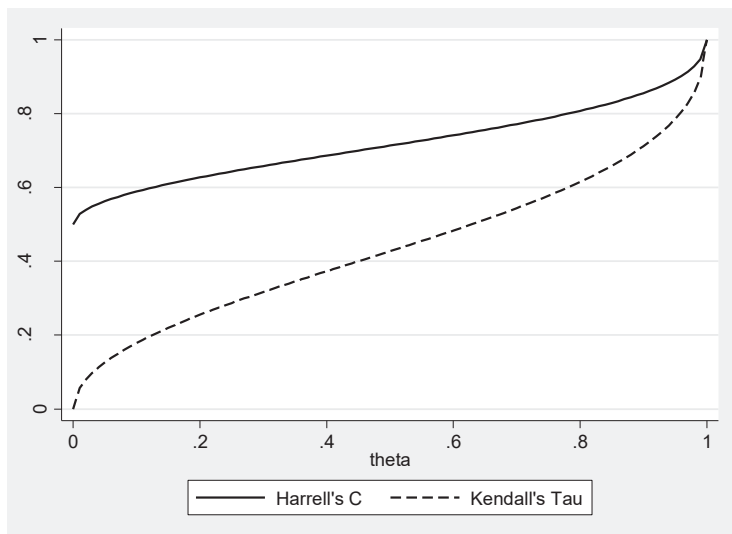


Рис. 1. Асимптотические коэффициенты Харрелла (сплошная линия) и Кендалла (пунктирная линия) для согласованности истинных показателей неэффективности и их оценок в зависимости от доли дисперсии неэффективности в дисперсии случайной составляющей ( $\theta$ )

Выражения (7)–(10) могут быть использованы в прикладных исследованиях для измерения точности ранжировок, полученных из



простой модели стохастической границы с экспоненциально распределённой компонентой неэффективности. При этом вместо истинных значений  $\rho$  и  $\theta$  нужно подставлять значения, рассчитанные на основании оценённых дисперсий  $\sigma_v^2$  и  $\sigma_u^2$ .

### Сравнение асимптотических коэффициентов корреляции со значениями, полученными на конечных выборках

В работе [14] приводятся результаты численного эксперимента, в котором коэффициент Харрелла рассчитывался для генерируемых выборок конечного объёма при разных соотношениях дисперсий неэффективности и случайного шока. В таблице 1 мы приводим некоторые из этих результатов (столбцы с указанием объёма выборок:  $n = 25$ ,  $n = 100$ ,  $n = 10000$ ) для сопоставления с асимптотическими значениями, рассчитанными по формуле (8).

Табл. 1. Коэффициенты Харрелла между истинными показателями неэффективности и их оценками в зависимости от объёма выборки. Источники: статья [14] и расчёты авторов

$\theta$	$n = 25$	$n = 100$	$n = 10000$	асимптотический
0.1	0.593	0.591	0.590	0.590
0.2	0.624	0.626	0.628	0.628
0.3	0.654	0.658	0.659	0.659
0.4	0.685	0.684	0.687	0.687
0.5	0.703	0.711	0.714	0.714
0.6	0.727	0.739	0.742	0.742
0.7	0.757	0.770	0.772	0.772
0.8	0.789	0.805	0.808	0.808
0.9	0.835	0.852	0.856	0.856

Как видно из таблицы, расчёты по представленной в настоящей статье формуле полностью согласуются с результатами исследования по искусственным данным. С практической точки зрения важно, что асимптотические значения близки к тем, что получены для малых ( $n = 25$ ) и умеренно больших ( $n = 100$ ) выборок.

### Пример приложения

Чтобы продемонстрировать практическую значимость полученных выражений, мы взяли результаты оценивания модели стохастической границы для отраслей французской промышленности из статьи [2], где среди прочего представлены и оценённые дисперсии стохастического шока и компоненты неэффективности. По этим оценкам были рассчитаны

асимптотические коэффициенты Харрелла и Кендалла, отражающие согласованность истинных показателей неэффективности предприятий и их оценок — результаты приведены в табл. 2.

Табл. 2. Асимптотические коэффициенты Харрелла ( $C_{\hat{a},u}^{asy}$ ) и Кендалла ( $\tau_{\hat{a},u}^{asy}$ ) между истинными показателями неэффективности предприятий и их оценками по данным об отраслях промышленности Франции из работы [2]

Отрасль	$\sigma_v^2$	$\sigma_u^2$	$\theta$	$C_{\hat{a},u}^{asy}$	$\tau_{\hat{a},u}^{asy}$
Изделия из стекла	0.7898	0.0036	0.0045	0.5190	0.0380
Молочные продукты	0.6438	0.0063	0.0097	0.5278	0.0555
Текстильная промышленность	0.9309	0.0100	0.0106	0.5291	0.0582
Машиностроение	0.7067	0.0114	0.0159	0.5355	0.0711
Электротехника	0.8154	0.0266	0.0316	0.5502	0.1003
Транспортные средства	0.7365	0.0451	0.0577	0.5678	0.1357
Промышленная химия	0.9307	0.0706	0.0705	0.5750	0.1501
Бумажная промышленность	0.6970	0.0717	0.0933	0.5864	0.1728
Обувь	0.3714	0.1042	0.2191	0.6340	0.2679
Сахар и напитки	0.9115	0.1719	0.1587	0.6133	0.2265

Из таблицы видно, что связь между оценками и истинными показателями практически отсутствует — наибольшая теснота наблюдается для предприятий обувной промышленности, где коэффициент Харрелла достигает 0.634. Таким образом, при сравнении пары предприятий с разным уровнем неэффективности, модель верно выделяет более эффективное предприятие с вероятностью лишь 0.634. Для многих отраслей эта вероятность близка к 0.5 — в таких случаях сравнение предприятий по регрессионной модели даёт примерно такую же точность, как и выделение относительно эффективного предприятия подбрасыванием монетки.

Рассчитанный коэффициент Кендалла (0.2679 для обувной промышленности) означает, что вероятность правильно выбрать более эффективное предприятие из двух на 0.2679 больше вероятности ошибиться.

В целом, можно сделать вывод, что модель не может дать точную ранжировку предприятий по рассматриваемым данным. Это ни в коей

мере не означает её непригодность для иных целей, помимо ранжирования — в самой статье [2] речь шла об измерении неэффективности отрасли в целом, оценки для отдельных предприятий на тот момент ещё не были выведены. Приведённые в табл. 2 числа также не означают, что модель стохастической границы вообще не способна давать точные ранжировки — согласованность истинных и оценённых показателей неэффективности будет высокой, если разброс случайных отклонений от линии регрессии будет связан в основном с неэффективностью, а дисперсия случайных шоков будет относительно низка.

### Заключение

Представленные в настоящей статье результаты могут помочь исследователям и регулирующим органам оценить применимость модели стохастической границы для ранжирования субъектов принятия решения. Как следует из формул (7)–(10), высокая точность ранжировок может достигаться только в тех случаях, когда дисперсия стохастических шоков низка по сравнению с различиями субъектов по уровню неэффективности. Приведённый в статье пример оценивания коэффициентов согласованности ранжировок по реальным данным показывает, что на практике это условие вполне может не выполняться.

Помимо отрицательной роли стохастических шоков есть иная причина для опасений: выполнение предпосылок модели стохастической границы. В первую очередь — предпосылки о распределениях шоков и показателей неэффективности. К сожалению, эта предпосылка на практике не проверяема, так как наблюдению поддаётся только суммарная ошибка регрессии, включающая как шок, так и неэффективность. Читатель, заинтересованный в последствиях неверного выбора закона распределения, может обратиться к статьям [25], [26].

### Литература

1. Айвазян С.А., Афанасьев М.Ю., Кудров А.В. Модели производственного потенциала и оценки технологической эффективности регионов РФ с учётом структуры производства // Экономика и математические методы, 2016, 52, 1, с. 28–44.
2. Айвазян С.А., Афанасьев М.Ю., Кудров А.В. Индикаторы экономического развития в базисе характеристик региональной дифференциации // Прикладная эконометрика, 2018, 50, с.4–22.
3. Головань С.В. Факторы, влияющие на эффективность российских банков // Прикладная эконометрика, 2006, 2, с.3–17.

4. Головань С.В., Карминский А.М., Пересецкий А.А. Эффективность российских банков с точки зрения минимизации издержек, с учетом факторов риска // Экономика и математические методы, 2008, 44, 4, с.28–38.
5. Головань С.В., Назин В.В., Пересецкий А.А. Непараметрические оценки эффективности российских банков // Экономика и математические методы, 2010, 46, 3, с.43–57.
6. Ипатов И.Б. Динамика совокупной факторной производительности и ее компонентов на примере российской отрасли, производящей пластмассовые изделия // Прикладная эконометрика, 2015, 38, с.21–40.
7. Макаров В.Л., Айвазян С.А., Афанасьев М.Ю., Бахтизин А.Р., Нанаян А.М. Оценка эффективности регионов РФ с учетом интеллектуального капитала, характеристик готовности к инновациям, уровня благосостояния и качества жизни населения // Экономика региона, 2014, 4, с.9–30.
8. Малахов Д.И., Пильник Н.П. Методы оценки показателя эффективности в моделях стохастической производственной границы // Экономический журнал Высшей Школы Экономики, 2013, 17 (4), с.660–686.
9. Матеров И.С. К проблеме полной идентификации модели стохастических границ производства // Экономика и математические методы, 1981, 17, 4, с.784–788.
10. Никольский И.М., Фурманов К.К. Измерение точности ранжировок предприятий по эффективности в модели стохастической границы // Прикладная эконометрика, 2023, 71, с.128-142.
11. Промахина И.М. Об одной характеристике востребованности жилищных кредитов в регионах РФ // В книге: Управление развитием крупномасштабных систем MLSD'2019. Материалы двенадцатой международной конференции. Под общей ред. С.Н. Васильева, А.Д. Цвиркуна. 2019, с.537-539
12. Щетинин Е.И., Назруллаева Е.Ю. Производственный процесс в пищевой промышленности: взаимосвязь инвестиций в основной капитал и технической эффективности. // Прикладная эконометрика, 2012, 28, с.63–84.
13. Aigner D., Lovell C.A.K., Schmidt P. (1977). Formulation and estimation of stochastic frontier function models // Journal of Econometrics, 6, pp.21–37.
14. Aleskerov F.T., Belousova V., Petrushchenko S. Models of Data Envelopment Analysis and Stochastic Frontier Analysis in the Efficiency

- Assessment of Universities //Automation and Remote Control, 2017, 78, 5, pp.902-923.
15. *Battese G.E., Coelli T.J.* Prediction of Firm-level Technical Efficiencies: With a Generalized Frontier Production Function and Panel Data //Journal of Econometrics, 1988, 38, pp.387–399.
  16. *Grushka Eli.* Characterization of exponentially modified Gaussian peaks in chromatography //Anal. Chem. 1972, 44, 11, pp.1733–1738
  17. *Harrell F.E., Califf R.M., Pryor D.B., Lee K.L. and Rosati R.A.* Evaluating the yield of medical tests //Journal of the American Medical Association, 1982, 247, pp.2543–2546.
  18. *Horrace W.C., Seth R.-S., Wright I.* Expected efficiency ranks from parametric stochastic frontier models. //Empirical Economics, 2015, 48, 2, pp.829–848.
  19. *Jondrow J., Lovell C.A.K., Materov I.S., Schmidt P.* On the estimation of technical inefficiency in stochastic frontier production function model //Journal of Econometrics, 1982, 19, pp.233–239.
  20. *Kirjavainen T.* Efficiency of Finnish upper secondary schools: an application of stochastic frontier analysis with panel data // Educ. Econ. 20(4), 2012, pp.343–364.
  21. *Kumbhakar S.C., Peresetsky A., Shchetynin Y., Zaytsev A.* Technical efficiency and inefficiency: Reliability of standard SFA models and a misspecification problem //Econometrics and Statistics (in press), 2021.
  22. *Meeusen W., Broeck van den J.* Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error //International Economic Review, 1977, 18(2), pp.435–444.
  23. *Newson R.* Parameters behind “nonparametric” statistics: Kendall’s tau, Somers’ D and median differences //Stata Journal, 2002, 2, pp.45–64.
  24. *Nikol’skii I.M., Furmanov K.K.* Ranking Accuracy of the Efficiency Index in the Stochastic Frontier Model //Computational Mathematics and Modeling, 2022, 33, 3, pp.319–329.
  25. *Sakouvogui K., Shaik S., Doetkott C., Magel R.* Sensitivity analysis of stochastic frontier analysis models //Monte Carlo Methods and Applications, 2021, 27, 1, pp.71–90.
  26. *Somers R.H.* A new asymmetric measure of association for ordinal variables //American Sociological Review, 1962, 27, pp.799–811.