УДК 519.2:330.54

СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ ПАРАМЕТРА КРИВОЙ ЛОРЕНЦА

к.т.н. В.Ю. Дубницкий, А.И. Ходырев
(представил д.ф.-м.н., проф. С.В. Смеляков)

Получены оценки параметра формы кривой Лоренца методом моментов, методом максимума правдоподобия, методом наименьших квадратов. Для оценки, полученной методом максимума правдоподобия, найдена величина доверительного интервала.

Распределение **F(x)** населения некоторого региона или государства в целом по уровню дохода принято называть кривой Лоренца. При ее построении по оси абсцисс откладывают численность отдельных групп населения, а по оси ординат уровень их дохода в порядке возрастания. Если провести нормировку по осям абсцисс и ординат, то построение будет ограничено единичным квадратом, а возможный вид получившейся кривой показан на рис.1.

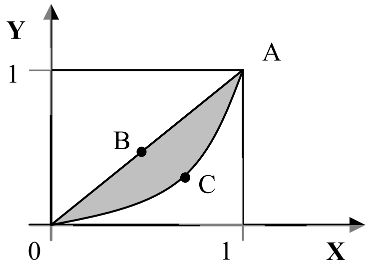

Рис.1. Общий вид кривой Лоренца

Здесь **Y** – уровень дохода в долях от максимального, **X** – доля населения, имеющего доход, не превосходящий **Y**; по построению $\mathbf{0 \le y < 1}$ и $\mathbf{0 \le x \le 1}$.

Кривая ОСА и есть собственно кривая Лоренца, а удвоенную площадь фигуры ОВАСО, величину $\mathbf{S_{OBACO}}$, называют коэффициентом Джини. Из построения фигуры ОВАСО следует, что чем она больше, тем неравенство в распределении населения по уровню доходов выше.

Обоснование этого приема с позиций экономической теории приведено в работах [1 - 3]. Основные результаты этого исследования приведены в настоящем сообщении.

Из способа построения кривой Лоренца $\mathbf{y = F(x)}$ и рисунка видно, что функция **F(x)** не убывающая, непрерывная слева, при $\mathbf{x \le 0}$ функция $\mathbf{F(x) = 0}$; при $\mathbf{x \ge 1}$ функция $\mathbf{F(x) = 1}$, т.е. соответствует требованиям, предъявляемым к функциям распределения.

Предположим, что

$$\mathbf{F(x) = x^c}, \quad \mathbf{c \ge 1}, \tag{1}$$

т.е. $\mathbf{F(x)}$ – степенное распределение.

Определение параметра $\mathbf{c}$ и является целью данного сообщения.

Плотность распределения вида (1) будет

$$\mathbf{f(x) = cx^{c-1}}. \tag{2}$$

Математическое ожидание $\mathbf{m_x}$ данного распределения

$$\mathbf{m_x = \int_0^1 cx^{c-1}dx = \frac{c}{c+1}}. \tag{3}$$

Используя для оценки величины $\mathbf{c}$ метод моментов получим, что оценка $\hat{\mathbf{c}}_{\mathbf{m}}$ величины $\mathbf{c}$ будет

$$\mathbf{\widehat{C}_m = \frac{m_x}{1-m_x}} \tag{4}$$

или

$$\mathbf{\widehat{C}_m = \frac{\frac{1}{n}\sum_{I-1}^{n} x_I}{1-\frac{1}{n}\sum_{I-1}^{n} x_I}}. \tag{5}$$

Функция правдоподобия примет вид

$$\mathbf{L(x;c) = C^n \prod_{I-1}^{n} x_I^{c-1}}, \tag{6}$$

а ее логарифм

$$\mathbf{\mathit{l}nL(x;c) = n\mathit{l}nC + (C-1)\sum_{I-1}^{n} \mathit{l}nx_I}. \tag{7}$$

Так как

$$\mathbf{\frac{d\mathit{l}nL}{dc} = \frac{n}{c} + \sum_{I-1}^{n} \mathit{l}nx_I} \tag{8}$$

и

$$\mathbf{\frac{d^2\mathit{l}nc}{dc} = -\frac{n}{c^2} < 0}, \tag{9}$$

то приравнивая (8) нулю и учитывая, что

$$\mathbf{\sum_{I=1}^{n} \mathit{l}nx_I = n\overline{\mathit{l}nx}},$$

получаем оценку $\mathbf{\widehat{C}_t}$ величины $\mathbf{C}$:

$$\hat{C}_t = \left(-\overline{l\mathrm{n}x}\right)^{-1} = \left(-\frac{1}{n}\sum_{I=1}^{n} l\mathrm{n}x_I\right)^{-1}. \tag{10}$$

Следуя формуле 6.2.15 из [4], можно утверждать, что оценка $\hat{C}_t$ параметра **C** распределена приближенно нормально с нулевым математическим ожиданием и дисперсией

$$\tau^2(c) = -E\left[\frac{\partial^2 l\mathrm{n}L}{dC^2}\right], \tag{11}$$

где **E [·]** – оператор математического ожидания, примененный к выражению, заключенному в квадратные скобки.

Используя (8) и (11) получим, что

$$\tau^2(c) = \frac{n}{c^2}. \tag{12}$$

Используя [4, стр. 302] получим, что приближенный 95% доверительный интервал для параметра **с** составит

$$\hat{C}_t - \frac{\hat{C}_t}{\sqrt{n}}\ C \le \hat{C}_t + \frac{2\hat{C}_t}{\sqrt{n}}. \tag{13}$$

В том случае, когда применение критерия χ^2 отвергает гипотезу о том, что кривая Лоренца совпадает с указанным в (1) степенным распределением, ее приближенное описание можно получить, используя метод наименьших квадратов, предложенный в [1. П]. Примем, что

$$y = ax^2 + (1-a)x, \tag{14}$$

где **a** – коэффициент, подлежащий определению.

Основное уравнение измерения

$$y_I = ax_I^2 + (1-a)x_I \tag{15}$$

или

$$y_I = a\left(x_I^2 - x_I\right) + x_I. \tag{16}$$

Если принять, что

$$y_I - x_I = U_I \quad \text{и} \quad x_I^2 - x_I = \omega_I,$$

то

$$U_i = a\omega_i.$$

Используя стандартную процедуру метода наименьших квадратов, получим оценку $\hat{a}_{LSq}$ параметра **a**:

$$\hat{a}_{LSq} = \frac{\sum_{I=1}^{n} U_I W_I}{\sum W_I^2}. \qquad (17)$$

Доверительные интервалы оценки $\hat{\mathbf{a}}_{LSq}$ могут быть получены по схеме, принятой в регрессионном анализе для аналогичных вычислений. Коэффициент Джини $\mathbf{G = 2S_{OBACO}}$ будет равен величине

$$G_d = 2\int_0^1 \left(x - x^c\right)dx = 1 - \frac{2}{c+1} \qquad (18)$$

в случае использования степенного распределения (1) и

$$G_A = 2\int_0^1 \left[x - \left(ax^2 + (1-\alpha)x\right)\right]dx = \frac{a}{3} \qquad (19)$$

в случае использования полиномиального приближения (14).

Предложенные в данной работе способы аппроксимации кривой Лоренца позволяют более строго с позиции математической статистики выполнять сравнительный анализ неравномерности доходов между регионами или исследовать ее в динамике.

ЛИТЕРАТУРА

1. Барковський В.В., Барковська Н.В. Математика для економістів. Вища математика. – К.: Національна академія управління, 1999. – 399 с.

2. Башарин Г.П. Начало финансовой математики. – М.: ИНФРА-М, 1997. – 160 с.

3. Сычев О.В. Оценка уровня доходов населения в условиях дифференциации общества // Системи обробки інформації. – Харків: НАНУ, ПАНМ, ХВУ. – 2001. – Вип. 3 (13). – С. 122 - 125.

4. Справочник по прикладной статистике. В 2-х т. – Т. 1.: пер. с англ. / Под ред. Э. Мойда, У. Ледермана, Ю. Н. Тюрина. – М.: Финансы и статистика. 1989. – 510 с.

Поступила в редколлегию 09.08.2001