Research article / Научная статья

Group averaging and the Gini deviation

Oleg I. Pavlov[1] ✉, Olga Yu. Pavlova[2]

[1]*Peoples' Friendship University of Russia (RUDN University),
6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*
[2]*All-Russian Correspondence Multidisciplinary School,
B-234 Vorob'evy Gory, Moscow, 119234, Russian Federation*
✉ pavlov_oi@rudn.ru

Abstract. It is known that partitioning a society into groups with subsequent averaging in each group decreases the Gini coefficient. The resulting Lorenz function is piecewise linear. This study deals with a natural question: by how much the Gini coefficient could decrease when passing to a piecewise linear Lorenz function? Obtained results are quite illustrative (since they are expressed in terms of the geometric parameters of the polygon Lorenz curve, such as the lengths of its segments and the angles between successive segments) upper bound estimates for the maximum possible change in the Gini coefficient with a restriction on the group shares, or on the difference between the averaged values of the attribute for consecutive groups. It is shown that there exist Lorenz curves with the Gini coefficient arbitrarily close to one, and at the same time with the Gini coefficient of the averaged society arbitrarily close to zero.

Keywords: Gini coefficient, Lorenz curve, Gini deviation

Усреднение по группам и изменение коэффициента Джини

О.И. Павлов[1] ✉, О.Ю. Павлова[2]

[1]*Российский университет дружбы народов,
Российская Федерация, 117198, Москва, ул. Миклухо-Маклая, д. 6*
[2]*Всероссийская заочная многопредметная школа,
Российская Федерация, 119234, Москва, Воробьевы горы, В-234*
✉ pavlov_oi@rudn.ru

Аннотация. Известно, что разбиение общества на группы с последующим усреднением в каждой группе приводит к уменьшению коэффициента Джини. Изучается вопрос, насколько может уменьшиться коэффициент Джини при переходе к данной кусочно-

линейной функции Лоренца. Получены весьма наглядные (так как они выражены через геометрические параметры графика кусочно-линейной функции Лоренца, такие как длины ее звеньев и углы между последовательными звеньями) оценки сверху на максимально возможное изменение коэффициента Джини при ограничении на величину доли групп либо на величину разности между усредненными значениями признака по данной группе и по предшествующей группе. Показано, что существуют кривые Лоренца с коэффициентом Джини, сколь угодно близким к единице, и при этом с коэффициентом Джини усредненного общества, сколь угодно близким к нулю.

Ключевые слова: коэффициент Джини, кривая Лоренца, отклонение коэффициента Джини, кусочно-линейная функция

Introduction:
the Gini coefficient of a piecewise linear Lorenz function

The Gini coefficient is one of the most important (income) inequality measures. Originally due to Gini (1912), also see (Ceriani, Verme, 2012), it is defined as twice the area of the region between the Lorenz curve and the diagonal OD (Figure 1). The Lorenz curve is the graph of the Lorenz function, that is, the function $L(x)$ defined on $[0; 1]$ such that $L(x)$ is equal to the fraction of the total income or wealth, owned by the poorest fraction x of the population. The Lorenz function is necessarily non-decreasing and concave up, i.e. $\frac{L(x) - L(x_0)}{x - x_0}$ is a non-decreasing function of x for every $x_0 \in [0; 1]$.
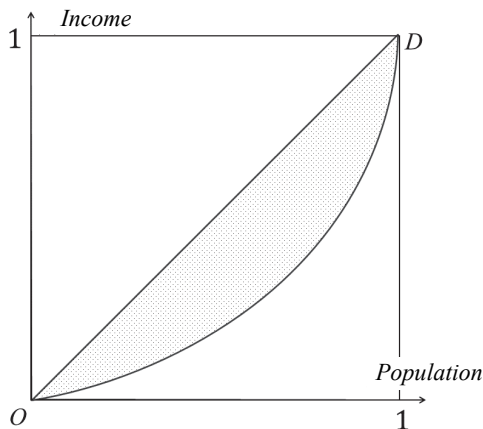
Figure 1. The Lorenz curve

Throughout the paper, the Gini coefficient corresponding to the Lorenz function $L(x)$ is denoted by $J(L)$ or simply by J. The Gini coefficient is widely used in economics and other fields of knowledge (see more details in: Arnold, 2007; Farris, 2010; Astashenko, Malykhin, 2012; Pavlov, Pavlova, 2016, 2018), but most often for measuring income inequality. For example, the Gini coefficient

for the total volume of monetary income of population in Russia in 2019 was 0,411.[1] In the European Union in 2019 the Gini coefficient for equivalised disposable income ranged from 0,228 for Slovakia to 0,408 for Bulgaria according to Eurostat.[2]

If the Lorenz curve is a polygonal chain (which corresponds to the division of society into groups with uniform distribution in each group) then the Gini coefficient of the society is (Kämpke, Radermacher, 2015):

$$J = 1 - \sum_{i=1}^{n}(x_i - x_{i-1})(y_i + y_{i-1}), \tag{1}$$

where $(x_0; y_0) = (0; 0)$, $(x_1; y_1)$, $(x_2; y_2)$,..., $(x_n; y_n) = (1; 1)$ are the vertices of this chain.

Consider a typical oval-shaped Lorenz curve L (with no segment of population having a uniform income distribution). Let $x_0 = 0 < x_1 < \cdots < x_{n-1} < x_n = 1$ be a finite increasing sequence on the x-axis. If we average out the income in each group $[x_0; x_1]$, $[x_1; x_2]$,..., $[x_{n-1}; x_n]$, then the resulting Lorenz curve L' will be piecewise linear and inscribed into original Lorenz curve L (Figure 2). Since L is concave up and does not contain a non-trivial segment, the only common ponts of L and L' are $(x_0; L(x_0))$, $(x_1; L(x_1))$,..., $(x_n; L(x_n))$. Hence L' is above L everywhere except these points, so the area of the region between L' and the diagonal of the square is strictly less than between L and this diagonal. Therefore $J(L') < J(L)$.
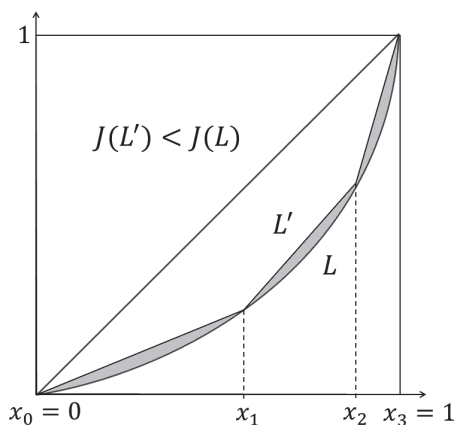
Figure 2. The original Lorenz curve L and its piecewise linearization L'

The procedure of partitioning the population into 5 or 10 equal groups (i.e. into quintiles or deciles) is a standard one and is used "for the convenience of computations and to increase the analyticity of the data."[3] Such data modification can lead to computational errors. How much does the Gini coefficient decrease when such a procedure is used?

---

[1] Federal State Statistics Service. (2020). *Russian Statistical Yearbook 2020* (p. 161). Moscow.

[2] Eurostat. (2021). *Gini coefficient of equivalised disposable income – EU-SILC survey*. Retrieved July 15, 2021, from http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=ilc_di12

[3] Minashkin, V.G. (2016). Gini index. *The Great Russian Encyclopedia. Electronic Version.* (In Russ.) Retrieved July 15, 2021, from https://bigenc.ru/economics/text/1952567

In the above mentioned total monetary income in Russia in 2019 the combined income of the five 20 percent groups is, respectively, 5,3; 10,1; 15,1; 22,6, and 46,9 percent of the total monetary income. It follows from (1) that this partition (with subsequent averaging in each group) corresponds to the Gini coefficient of 0,319, which is considerably less than 0,411!

Literature review

Estimation of the Gini coefficient from incomplete/grouped data has been studied extensively. Since grouped data leads to Lorenz curve which is a polygonal chain, and the definite integral over the latter is computed with the trapezoidal rule exactly, any estimation formula for this rule (e.g., see: Zorich, 2019) produces an estimate for the Gini coefficient $J(L)$. However, there are better alternatives. For example, Gastwirth (1972) presented several methods for calculation of the lower and upper bounds using a probability density function. Golden (2008) proposed an easy to compute estimate for data presented in cumulative income quintiles using a single point of the Lorenz curve. Farris (2012) gave interval estimates using aggregate parameters (such as the number of individuals/families in a percentile and their overall wealth). Fellman (2012) surveyed further publications and concluded that "no method is uniformly optimal, but the trapezium rule is almost always inferior, and Simpson's rule is superior. Golden's method is usually of medium quality."

In this note we obtain upper bound estimates for the maximum possible change in the Gini coefficient with a restriction on the group shares, or on the difference between the averaged values of the attribute for consecutive groups. These estimates are as illustrative as in (Golden, 2008) because they are expressed in terms of the geometric parameters of the polygonal Lorenz curve (such as the lengths of its segments and the angles between successive segments), but some of them are more accurate. Let L be an arbitrary Lorenz function/curve and P – some finite increasing sequence on the x-axis $x_0 = 0 < x_1 < \cdots < x_{n-1} < x_n = 1$. Then L_P or simply L' will denote the Lorenz function/curve which is obtained from L by linearization on segments whose endpoints are consecutive elements of P. We will call L' a piecewise inearization of L corresponding to P and call the difference $J(L) - J(L')$ *the deviation* of the Gini coefficient or simply the Gini deviation. We show that the deviation of the Gini coefficient can be arbitrary close to 1 (i.e., similtaniously $J(L)$ can be arbitrary close to 1, and $J(L')$ to 0).

A Lorenz curve with a large Gini deviation

Example 1. Consider a Lorenz curve L which is a polygonal chain $OABCD$, where $O = (0; 0)$, $D = (1; 1)$, $A = (\varepsilon; \varepsilon^2)$, $C = (1 - \varepsilon^2; 1 - \varepsilon)$ and ε is small (so that A is much closer to the x-axis than to the y-axis and C is much closer to the vertical line $x = 1$ than to the horizontal line $y = 1$, see Figure 3).

The lines OA and CD have equations $y = \varepsilon x$ and $y = \frac{1}{\varepsilon} x + \frac{\varepsilon - 1}{\varepsilon}$ respectively; we denote their intersection by B. By solving the system of these equations, we get $B = \left(\frac{1}{1+\varepsilon}; \frac{\varepsilon}{1+\varepsilon} \right)$. As B is close to the vertex $(1; 0)$, the area of the region between L and the diagonal OD is close to the area of the triangle with the vertices

O, D and $(1; 0)$, and the latter area equals 1/2. This means that $J(L)$ is close to 1. $J(L)$ could be calculated exactly using (1), but it is easy to do so by doubling the difference between 1/2 and the area of the region under L. This region is a union of a rectangle (actually, a square) and two triangles. Its area is equal to

$$\frac{1}{2} \cdot \frac{1}{1+\varepsilon} \cdot \frac{\varepsilon}{1+\varepsilon} + \frac{1}{2}\left(1 - \frac{1}{1+\varepsilon}\right)\left(1 - \frac{\varepsilon}{1+\varepsilon}\right) + \left(1 - \frac{1}{1+\varepsilon}\right)\frac{\varepsilon}{1+\varepsilon} = \frac{\varepsilon + \varepsilon^2}{1 + \varepsilon^2}.$$

$$J(L) = 2\left(\frac{1}{2} - \frac{\varepsilon + \varepsilon^2}{1 + \varepsilon^2}\right) = 1 - \frac{2\varepsilon + 2\varepsilon^2}{1 + \varepsilon^2}.$$
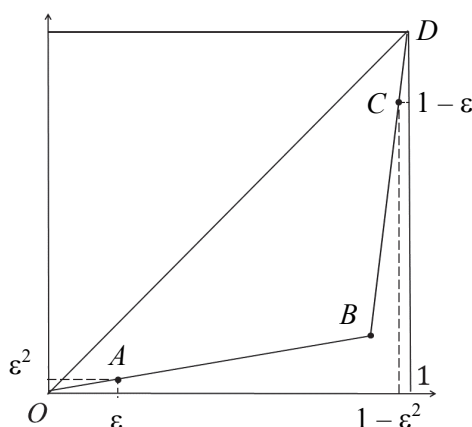
Thus, $J(L)$ is indeed close to 1 for small ε.

Figure 3. Lorenz curve L

Let $P = \{0; \varepsilon; 1 - \varepsilon^2; 1\}$, then L_P is a polygonal chain $OACD$. Since the points A and C are close to the diagonal, $J(L_P)$ is close to 0. We calculate it using (1). The differences $x_i - x_{i-1}$ equal $\varepsilon - 0$, $1 - \varepsilon^2 - \varepsilon$, $1 - \varepsilon^2$ for $i = 1$, 2, 3 respectively; the sums $y_i + y_{i-1}$ equal $\varepsilon^2 + 0$, $1 - \varepsilon + \varepsilon^2$, $2 - \varepsilon$ respectively.

$$J(L_P) = 1 - [\varepsilon \cdot \varepsilon^2 + (1 - \varepsilon^2 - \varepsilon)(1 - \varepsilon + \varepsilon^2) + \varepsilon^2(2 - \varepsilon)] = 2\varepsilon - 3\varepsilon^2 + \varepsilon^4.$$

Therefore, both $J(L)$ and $1 - J(L_P)$ are of order $O(\varepsilon)$ for the considered Lorenz curve L. This means that $J(L)$ is close to 1, and $J(L_P)$ is close to 0 for a small ε.

Note the following features of the considered L: the share of one of the groups (the second one) is close to 1 and the angle between some segments of L (namely, between AB and BC and also between AB and CD) is close to 90°.

These features are not accidental. We show below that both the smallness of the share of each group or the smallness of the angle between every two consecutive segments of L (or of every two segments standing the next but one) implies the smallness of the Gini deviation.
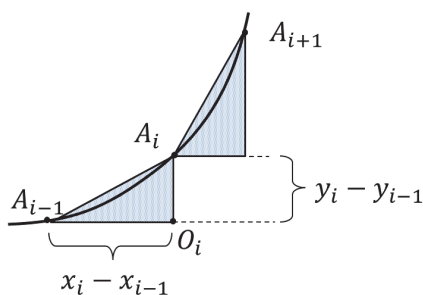
The Gini deviation and the shares of groups

Theorem 1. Let d be the largest share of groups in a society (i.e. the maximum of values $x_i - x_{i-1}$ for $1 \le i \le n$). Then $J(L) - J(L') \le d$.

Proof. Let L' be a piecewise linearization of the Lorenz curve L corresponding to a sequence $P = x_0 = 0 < x_1 < \cdots < x_{n-1} < x_n = 1$. For every $i \le n$, y_i will denote the common value of $L(x_i)$ and $L'(x_i)$.

If $n = 1$, then the society is "divided" into a single group whose share is equal to 1. The theorem is valid in this case as the Gini deviation never exceeds 1. So we can assume that $n > 1$.

Consider two consecutive vertices $A_{i-1} = (x_{i-1}; y_{i-1})$ and $A_i = (x_i; y_i)$ of L'. Since L is a graph of a non-decreasing function, its entire piece lying between A_{i-1} and A_i is contained in the right triangle $A_{i-1}A_iO_i$, which is bounded by the hypotenuse $A_{i-1}A_i$ from above and by a horizontal leg $A_{i-1}O_i$ and a vertical leg A_iO_i from below (Figure 4).

Figure 4. The Gini deviation and the shares of groups

Consequently, the part of the region between L' and L which lies between A_{i-1} and A_i is also included in this triangle. Therefore, the whole region between L' and L is included in the union of such right triangles. Hence it is enough to show that the combined area of these triangles does not exceed d.

In the triangle $A_{i-1}A_iO_i$, the leg $A_{i-1}O_i$ is equal to $x_i - x_{i-1}$ and the leg A_iO_i is equal to $y_i - y_{i-1}$, so its area is equal to $\frac{1}{2}(x_i - x_{i-1})(y_i - y_{i-1})$. The deviation $J(L) - J(L')$ is equal to twice the area of the region between L' and L, which does not exceed $\sum_{i=1}^{n}(x_i - x_{i-1})(y_i - y_{i-1})$ according to the previous calculations. But $\sum_{i=1}^{n}(x_i - x_{i-1})(y_i - y_{i-1}) \le \sum_{i=1}^{n} d(y_i - y_{i-1}) \le d \cdot 1 = d$. Theorem 1 is proved.

If a society is divided into n equal groups, then share each of them is equal to $\frac{1}{n}$. Then, according to Theorem 1, the Gini deviation does not exceed $1/n$. Therefore, the maximum possible Gini deviation in the above mentioned total monetary income distribution in Russia in 2019 is $1/5 = 0,2$. The real deviation for this distribution is equal to $0,411 - 0,319 = 0,092$, which is a significant fraction of $0,2$. We obtained $J(L') = 0,319$ using (1) and the data from Federal State Statistics Service.[4]

---

[4] Federal State Statistics Service. (2020). *Russian Statistical Yearbook 2020* (p. 161). Moscow.

The Gini deviation estimate given by Theorem 1 is sharp for every d. For example, in case of perfect inequality in a society (all the income/wealth belongs to a single person), $J = 1$. The corresponding Lorenz curve L is a two-segment polygonal chain ABD, where B almost coincides with $(0; 1)$. Divide such a society into several groups so that the share of the richest of them is equal to an arbitrary d. Then L' will consist of a horizontal segment OB', where $B' = \left(\frac{n-1}{n}; 0\right)$ (this segment may correspond to several groups of the society) and an inclined line segment $B'D$. The doubled area between L' and the diagonal OD is equal to $1 - d$ which is $J(L')$. Thus, $J(L) - J(L') = 1 - (1 - d) = d$ as required.

The Gini deviation and the angles between segments

Every two consecutive segments of a polygonal Lorenz curve are at the angle from 0 to 90° to each other. The more segments, the smaller this angle can be uniformly given a fixed Gini coefficient. It turns out that a constraint on the magnitude of such an angle also limits the Gini deviation. It is also possible to consider a constraint on segments standing the next but one. Since the slope of the segment[5] with the vertices $(x_i; y_i)$ и $(x_{i+1}; y_{i+1})$ is the average income in the group $[x_i; x_{i+1}]$, the proximity of the slopes of two consecutive groups is equivalent to the closeness of the average incomes in these groups. We denote the points $(-1; 0)$ and $(1; 2)$ by A_{-1} and A_{n+1}, respectively, and will consider $A_{-1}A_0$ and DA_{n+1} as new segments of L'.

Theorem 2. Let α be the largest of the angles between the segments of L' standing the next but one. Then

a) if $l_1, l_2,..., l_n$ denote the lengths of the segments $A_0A_1, A_1A_2,..., A_{n-1}A_n$, then $J(L) - J(L') \leq \frac{1}{4}\sum_{i=1}^{n}(l_i)^2 \,\text{tg}\left(\frac{\alpha}{2}\right)$;

b) if $p_1, p_2,..., p_n$ denote the lengths of the projections of the segments A_0A_1, $A_1A_2,...,A_{n-1}A_n$ on the diagonal OD, then $J(L) - J(L') \leq \frac{1}{2}\sum_{i=1}^{n}(p_i)^2 \,\text{tg}\left(\frac{\alpha}{2}\right)$;

c) if $\frac{1}{n}$ is the largest fraction of the projections of the segments of L' on the diagonal OD, then $J(L) - J(L') \leq \frac{1}{n}\,\text{tg}\left(\frac{\alpha}{2}\right)$;

d) $J(L) - J(L') \leq \text{tg}\left(\frac{\alpha}{2}\right)$.

In each item **a–d**, $\frac{\alpha}{2}$ can be replaced by the largest of the angles between any two *consecutive* segments of L'.

Proof:

a) we use the following well-known fact from geometry (e.g. see: Boltyanskij et. al., 1974): if the length of a side S of a triangle is equal to l, and the other two sides make angles with S that in total do not exceed α, then the area of such a triangle does not exceed $\frac{1}{4}l^2\text{tg}\left(\frac{\alpha}{2}\right)$. Consider an arbitrary segment $A_{i-1}A_i$ of L', as well as the previous and subsequent segments (Figure 5).

---

[5] More precisely, of a straight line containing this segment.
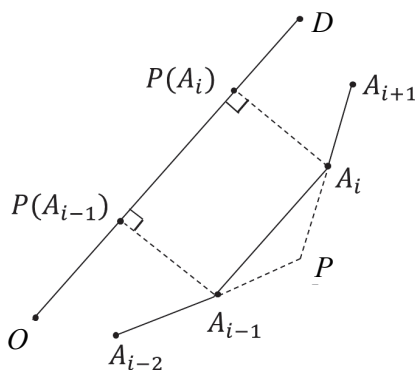
Figure 5. The Gini deviation and the angles between segments

Due to concavity of L, every point lying between the arc $A_{i-1}A_i$ of L (the arc is not shown in Figure 5) and the segment $A_{i-1}A_i$, must be not lower than both lines $A_{i-2}A_{i-1}$ and A_iA_{i+1}, i.e. it belongs to triangle $A_{i-1}A_iP$, where P is the intersection of the lines $A_{i-2}A_{i-1}$ and A_iA_{i+1}. By condition, the angle between the lines $A_{i-2}A_{i-1}$ and A_iA_{i+1} is at most $\propto$. Then, by the mentioned fact, the area of the triangle $A_{i-1}A_iP$ does not exceed

$$\frac{1}{4}|A_{i-1}A_i|^2 \operatorname{tg}\left(\frac{\propto}{2}\right) = \frac{1}{4}(l_i)^2 \operatorname{tg}\left(\frac{\propto}{2}\right).$$

If the lines $A_{i-2}A_{i-1}$ and A_iA_{i+1} coincide, then triangle $A_{i-1}A_iP$ degenerates into segment $A_{i-1}A_i$ having zero area. Therefore $J(L) - J(L') \le \frac{1}{4}\sum_{i=1}^{n}(l_i)^2 \operatorname{tg}\left(\frac{\propto}{2}\right)$. Item **a** is proved;

b) let $p(A_{i-1})$ and $p(A_i)$ denote, respectively, the projections of points A_{i-1} and A_i on the diagonal OD. Then the length of the segment $A_{i-1}A_i$ does not exceed the length of segment $p(A_{i-1})p(A_i)$ multiplied by $\sqrt{2}$. Thus $l_i \le p_i\sqrt{2}$ and further

$$J(L) - J(L') \le \frac{1}{4}\sum_{i=1}^{n}\left(p_i\sqrt{2}\right)^2 \operatorname{tg}\left(\frac{\propto}{2}\right) = \frac{1}{2}\sum_{i=1}^{n}(p_i)^2 \operatorname{tg}\left(\frac{\propto}{2}\right).$$

Item **b** is proved;

c) it follows directly from the condition of item **c** that $p_i \le \frac{1}{n}\sqrt{2}$ whence

$$J(L) - J(L') \le \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\sqrt{2}}{n}\right)^2 \operatorname{tg}\left(\frac{\propto}{2}\right) = \frac{1}{n^2}n \operatorname{tg}\left(\frac{\propto}{2}\right) = \frac{1}{n}\operatorname{tg}\left(\frac{\propto}{2}\right);$$

d) follows from **c** as $\operatorname{tg}\left(\frac{\propto}{2}\right) \ge \frac{1}{n}\operatorname{tg}\left(\frac{\propto}{2}\right)$ since $n \ge 1$.

If β is the largest of the angles between any two consecutive segments of L', then the largest of the angles between the segments of L' standing the next but one does not exceed 2β, so $\frac{\propto}{2}$ can be replaced by β in each item **a–d**. Theorem 2 is proved.

A comparison with Golden's approach

Golden (2008) proposed a simple way to compute the upper bound Gini co-efficient estimate for data presented in cumulative income quintiles. We compare the Gini deviation corresponding to this estimate, obtained for a particular Lorenz function $L(x) = x^2$, with the estimates from our Theorem 2 and find that the estimates given in items a, b, c of Theorem 2 are more accurate.

Golden called his method for approximating the upper bound of the Lorenz curve's L Gini coefficient "the Z-gradient rule". The estimate given by this rule is the Gini coefficient of a three-segment polygonal chain $OPQD$ (which we further denote by L_G), where $O = (0; 0)$, $D = (1; 1)$, and points P and Q lie on the x-axis and y-axis, respectively, so that the line PQ is tangent to L at a point $(z_0, L(z_0))$ of L farthest from the diagonal OD. It is known that $L'(z_0) = 1$, see Hoover (1936) and Kakwani (1980). Since $L(x) = x^2$, $L'(x) = 2x$, so that $z_0 = 0,5$ and $L(z_0) = 0,25$. The tangent line to the Lorenz curve at the point $(0,5; 0,25)$ has an equation

$$y - 0,25 = 1 \cdot (x - 0,5),$$

or $y = x - 0,25$, thus $P = (0,25; 0)$ and $Q = (1; 0,75)$. Therefore, $J(L_G) = \frac{7}{16}$ by (1). Let $x_0 = 0$, $x_1 = 0,2$, $x_2 = 0,4,\ldots$, $x_5 = 1$ be the first coordinates of the vertices of L' (which is the linearization of $L(x) = x^2$ corresponding to the division of the segment $[0; 1]$ into five equal subsegments). Then $J(L') = 0,32$ by (1), therefore

$$J(L_G) - J(L') = \frac{7}{16} - 0,32 = 0,1175,$$

which is the error of Golden's approximation.

We organize the calculations of the estimates from Theorem 2 in the Table.

Here k_i is the slope of the i-th segment of the linearized Lorenz curve L' and ϕ_i is its angle with the positive direction of the x-axis (measured in radians). The largest half-angle $\alpha/2$ between the segments of L' standing the next but one is $0,294$ rad (the angle between the first and the third segments), its tangent is equal to $0,303$. The sum of the squares of the segments of L' is equal to $0,464$, the sum of the squares of the projections of these segments on the diagonal OD is $0,432$.

Calculations of the estimates from Theorem 2

x_i	y_i	Δx_i	Δy_i	$k_i = \Delta y_i/\Delta x_i$	ϕ_i	$(\phi_i - \phi_{t-2})/2$	$\phi_i - \phi_{t-1}$	l_i^2	p_i^2
0	0								
0,2	0,04	0,2	0,04	0,2	0,197			0,042	0,029
0,4	0,16	0,2	0,12	0,6	0,54		0,343	0,054	0,051
0,6	0,36	0,2	0,2	1	0,785	0,294	0,245	0,08	0,08
0,8	0,64	0,2	0,28	1,4	0,951	0,205	0,165	0,118	0,115
1	1	0,2	0,36	1,8	1,064	0,139	0,113	0,17	0,157
Σ								0,464	0,432
max						0,294	0,343		

Consequently, the estimates obtained in items *a*, *b*, *c* of Theorem 2 are

$$\frac{1}{4}\sum_{i=1}^{5}(l_i)^2 \, \text{tg}\left(\frac{\propto}{2}\right) = 0,035,$$

$$\frac{1}{2}\sum_{i=1}^{5}(p_i)^2 \, \text{tg}\left(\frac{\propto}{2}\right) = 0,065,$$

$$\frac{1}{5}\,\text{tg}\left(\frac{\propto}{2}\right) = 0,061$$

respectively (rounded to the nearest thousandth), which are considerably more accurate, than 0,1175 obtained by applying the Golden's Z-gradient rule. The largest angle β between any two consecutive segments of L' is 0,343 rad (the angle between the first and the second segments). Its tangent is equal to 0,357, which is 18% greater, than $\text{tg}\,\alpha = 0,303$. Hence replacing $\propto/2$ with β in the latter three displayed formulas increases the estimates by 18% which still leaves them less than 0,1175.

Conclusion

We obtained upper bound estimates for the maximum possible change in the Gini coefficient with a restriction on the group shares, or on the difference between the averaged values of the attribute for consecutive groups (geometrically – in terms of the geometric parameters of the polygon Lorenz curve, such as the lengths of its segments and the angles between successive segments). For a considered particular Lorenz curve these estimates are much more accurate than derived by Golden (2008). Further research may include deriving still better (ultimately, sharp) estimates, e.g. such that for a given $\varepsilon > 0$ the Gini deviation could not be multiplied by $1 - \varepsilon$. The case of equal groups is of particular interest.

References

Arnold, B.C. (2007). The Lorenz curve: Evergreen after 100 years. In S. Betti, A. Lemmi (Eds.), *Advances in Income Inequality Concentration Measures* (pp. 12–24). New York: Routledge.

Astashenko, A.N., & Malykhin, V.I. (2012). *Income inequality measures*. LAP Lambert Academic Publishing.

Boltyanskij, V.G., Sidorov, Yu.V., & Shabunin, M.I. (1974). *Lectures and problems on elementary mathematics*. Moscow: Nauka Publ. (In Russ.)

Ceriani, L., & Verme, P. (2012). The origins of the Gini index: Extracts from Variabilità e Mutabilità (1912) by Corrado Gini. *J. Econ. Inequal, 10*, 412–443.

Farris, F.A. (2010). The Gini index and measures of inequality. *American Mathematical Monthly, 117*(10), 851–864.

Fellman, J. (2012). Estimation of Gini coefficients using Lorenz curves. *Journal of Statistical and Econometric Methods*, 1(2), 31–38.

Gastwirth, J. (1972). The estimation of the Lorenz curve and Gini index. *Rev. Econom. Statist, 54*, 306–316.

Gini, C. (1912). *Variabilità e mutuabilità: Contributo allo studio delle distribuzioni e delle relazioni statistiche*. Bologna: C. Cuppini.

Golden, J. (2008). A simple geometric approach to approximating the Gini coefficient. *The Journal of Economic Education, 39*(1), 68–77.

Hoover, E. (1936). The measurement of industrial localization. *The Review of Economics and Statistics, 18*, 162–171.

Kakwani, N. (1980). *Income inequality and poverty: Methods of estimation and poverty applications*. Oxford University Press.

Kämpke, T., & Radermacher, F.J. (2015). Income modeling and balancing. A rigorous treatment of distribution patterns. *Lecture Notes in Economics and Mathematical Systems, 679*, 44–53.

Pavlov, O.I., & Pavlova, O.Yu. (2016). The Lorenz curve and a mathematical definition of the middle class. *Management of Economic Systems*, (12). Retrieved March 15, 2021, from http://uecs.ru/uecs-94-942016/item/4239-2016-12-24-07-45-16

Pavlov, O.I., & Pavlova, O.Yu. (2018). Differential deviation and the Gini coefficient. *Russian Economics Online-Journal*, (4). Retrieved March 15, 2021, from http://www.e-rej.ru/publications/176/%D0%9F/

Zorich, V.A. (2019). *Mathematical analysis* (part 1). Moscow: MCCME Publ. (In Russ.)

Сведения об авторах / Bio notes

Павлов Олег Иванович, кандидат физико-математических наук, доцент кафедры экономико-математического моделирования, экономический факультет, Российский университет дружбы народов. E-mail: pavlov-oi@rudn.ru

Oleg I. Pavlov, PhD, Associate Professor of Economic and Mathematic Modelling Department, Economic Faculty, Peoples' Friendship University of Russia (RUDN University). E-mail: pavlov-oi@rudn.ru

Павлова Ольга Юрьевна, кандидат физико-математических наук, доцент кафедры высшей математики, Всероссийская заочная многопредметная школа. E-mail: lolgau@yandex.ru

Olga Yu. Pavlova, PhD, Associate Professor at the Department of Higher Mathematics, All-Russian Correspondence Multidisciplinary School. E-mail: lolgau@yandex.ru