

ЕЩЕ РАЗ О КОЭФФИЦИЕНТЕ ДЖИНИ КАК ПОКАЗАТЕЛЕ КОНЦЕНТРАЦИИ

**Г.Л. Громыко,
И.Н. Матюхина**

На страницах журнала в двух его номерах опубликованы две статьи, посвященные применению коэффициента Джини: «Об использовании коэффициента Джини в экономико-статистических исследованиях» (авторы: Г.Л. Громыко, И.Н. Матюхина) в № 9 за 2015 г. и «К вопросу о применении коэффициента Джини и других показателей неравенства» (автор: К.П. Глушенко) в № 2 за 2016 г. Авторы статей высказывают различные точки зрения на область применения широко известного коэффициента Джини.

Данная статья является продолжением развернувшейся дискуссии и содержит как ответ авторов первой статьи на критические замечания, изложенные во второй статье, так и критику основных ее положений. Различия в подходах авторов обнаруживаются по ряду вопросов: область применения коэффициента Джини как индекса концентрации, непосредственно связанного с кривой Лоренца и предназначенного для оценки неравномерности распределения исследуемого показателя по группам; трактовка понятий «распределение» и «концентрация» в статистике; выбор группировочного признака при построении вариационных рядов распределения; содержательная интерпретация показателей концентрации и дифференциации (различий) и др.

Ключевые слова: кривая Лоренца, коэффициент Джини, индекс концентрации, распределение населения по среднедушевым денежным доходам, распределения и группировки в статистике, средние и относительные величины, показатели дифференциации.

JEL: C10, C1.

В журнале «Вопросы статистики» (№ 9 за 2015 г.) была опубликована наша статья «Об использовании коэффициента Джини в экономико-статистических исследованиях», в конце которой прозвучал призыв к читателям принять участие в обсуждении некоторых, на наш взгляд, спорных вопросов, касающихся области применения коэффициента Джини как показателя (индекса) концентрации.

В ответ, в № 2 журнала за 2016 г. опубликована статья К.П. Глушенко «К вопросу о применении коэффициента Джини и других показателей неравенства», в которой, по словам самого автора, «излагается точка зрения по поводу применения средних, относительных и некоторых других величин для расчета коэффициента Джини» (2, с. 71, аннотация) и высказан ряд критических замечаний в адрес статьи (1)¹.

За отклик на нашу статью мы выражаем благодарность уважаемому Константину Павловичу, но наши взгляды по затронутому воп-

росу далеко не совпадают, и с большинством критических замечаний и утверждений оппонента мы не можем согласиться. Поэтому считаем своим долгом дать дополнительные разъяснения по некоторым вопросам и прежде всего определить предмет наших споров.

В первую очередь следует отметить, что хотя обе вышеупомянутые статьи посвящены коэффициенту Джини, речь в них по существу идет о разных вещах.

В нашей статье (1) коэффициент Джини рассматривается сугубо как индекс концентрации, связанный непосредственно с кривой Лоренца и характеризующий неравномерность распределения определенного исследуемого показателя по тем или иным группам единиц совокупности (например, предприятиям, регионам и т. п.). В статье приведены примеры расчета коэффициента Джини (G) для ряда показателей по различным исходным данным и примеры некорректного, с точки зрения авторов, использования (рас-

Громыко Галина Леонтьевна (gromyuko@econ.msu.ru) - д-р экон. наук, профессор кафедры статистики экономического факультета МГУ им. М.В. Ломоносова.

Матюхина Ирина Николаевна (iri1256@yandex.ru) - канд. экон. наук, доцент кафедры статистики экономического факультета МГУ им. М.В. Ломоносова.

¹ В дальнейшем тексте для упрощения изложения будем именовать их, не приводя названий статей, как статья (1) и статья (2) с учетом последовательности их публикации.

чета) коэффициента Джини как индекса концентрации применительно к среднелюдским показателям. В целом сделан вывод о том, что коэффициент Джини как индекс концентрации может использоваться лишь для объемных (суммарных) показателей, распределенных по группам с определением доли последних в общем объеме исследуемого показателя (признака). Использование же его для средних и относительных величин, не допускающих суммирования и определения кумулятивных долей в общем их итоге, бессодержательно. Не имеет смысла, например, рассчитывать сумму среднелюдских доходов населения регионов, а тем более, искать долю среднелюдского дохода населения каждого региона в этой сумме. Средние и относительные величины не аддитивны по своей природе.

В статье К.П. Глушенко, как видно из названия, речь идет о так называемых показателях неравенства, среди которых, по словам автора, «коэффициент Джини - лишь один из широкого класса измерителей неравенства, существо которых одно и то же» (2, с. 72). В статье (2) поясняется: «показатели неравенства - наиболее распространенный термин, но их можно назвать и показателями дифференциации, неравномерности, различий, разброса, концентрации, в зависимости от содержания изучаемого явления» (2, с. 73). Автор статьи (2) высказывает несогласие с нашим выводом об области применения коэффициента Джини как индекса концентрации и пытается доказать «правомерность использования коэффициента Джини и сходных показателей, рассчитанных на основе средних и относительных величин для измерения неравенства (дифференциации, неравномерности, различий и т. п.)» (2, с. 71, аннотация). В частности, для измерения межрегионального неравенства (различия) по таким показателям, как среднелюдской доход, уровень рождаемости, уровень преступности и др. При этом в качестве критических замечаний оппонент отмечает, что нами «неверно трактуется понятие распределения применительно к статистике» (2, с. 72) и абсолютизируется понимание коэффициента Джини как индекса концентрации и его связи с кривой Лоренца (2, с. 79).

Что можно сказать по поводу вышеуказанных замечаний? Они нас скорее удивляют,

чем огорчают. Трудно спорить и доказывать правоту или ошибочность тех или иных утверждений, когда четко не определен предмет спора и каждый говорит о своем.

1. На наш взгляд, понятие *концентрации* какого-либо показателя (явления) по группам единиц совокупности (например, регионов) и *различия* (неравенства) в значениях исследуемого показателя по группам (регионам) - далеко не одно и то же, и представляется весьма сомнительным, что один и тот же показатель может быть успешно использован для решения таких разных задач. Для оценки различий (неравенства) в статистической практике применяется ряд других показателей, в частности фондовый и децильный коэффициенты дифференциации или иные коэффициенты, построенные по квантилям; все они могут быть применены для измерения межрегионального неравенства по таким показателям, как среднелюдской доход, уровень рождаемости, уровень преступности и т. д.

2. О разных вещах идет речь в обеих статьях и тогда, когда мы говорим о некорректности расчета коэффициента Джини как индекса концентрации для средних и относительных показателей (среднелюдских, на 1000 человек и др.), то есть применительно к ним, а оппонент понимает это как расчет по средним величинам (на основании использования средних величин) и считает такое использование приемлемым. Мы же видим различия в понятиях «для какого показателя» и «по каким данным» рассчитывается тот или иной коэффициент. И между прочим, в статье (1) нами отмечалось, что при определенных условиях коэффициент Джини может быть рассчитан и по средним величинам, но для суммарного показателя; и это было проиллюстрировано на примере (в таблице 5).

3. Особо следует остановиться на критическом замечании нашего оппонента, что в статье (1) «неверно трактуется понятие распределения применительно к статистике. Из-за этого и ошибочный вывод об области применения коэффициента Джини», и что мы понимаем «распределение в быденном смысле» (2, с. 72).

Автор уточняет: «статистика-то пользуется иным понятием распределения, подразу-

мевающим эмпирическое распределение вероятностей» (2, с. 72) и его можно построить для любой совокупности, состоящей из единиц $i = 1, 2, 3, \dots, n$, каждая из которых имеет некоторую числовую характеристику значения признака. И что природа признака не имеет никакого значения. Говоря иначе, речь идет о построении вариационных рядов, где в результате группировки единиц совокупности по тому или иному признаку получается ряд распределения единиц совокупности. И действительно, природа группировочного признака не имеет никакого значения. Но никто с этим и не спорит. И в своей статье (1) мы демонстрировали расчет коэффициента Джини на примере вариационных рядов. Непонятно только, почему оппонент считает, что термин (слово) «распределение» не может использоваться за пределами понятия вариационного ряда (в «обыденном смысле», как говорит оппонент) и почему последнее следует оценивать как неверную трактовку распределения «применительно к статистике».

Говоря о распределении, мы различаем, как и должно: 1) распределение единиц совокупности по тому или иному количественному признаку (например, распределение населения по среднедушевому денежному доходу или по возрастным группам, распределение работников предприятий по стажу работы или по размеру заработной платы и др.) и 2) распределение общего объема определенного количественного показателя по группам единиц, в том числе и по регионам (например, распределение общего объема денежных доходов населения по 20%-ным группам населения, распределение производства отдельных видов сельскохозяйственной продукции (или поголовья скота) по регионам и т. п. В первом случае речь идет о вариационных рядах распределения. Во втором случае - тоже о распределении, но по форме оно не является вариационным рядом. Другими словами, не каждое распределение по форме является вариационным рядом.

Так, например, для данных, публикуемых в статистических справочниках в региональном разрезе, вполне возможно использование термина «распределение», но избирательно: для одних показателей (объемных) его можно применять, а для других - нет. Например,

можно говорить о распределении по регионам численности населения, числа безработных, поголовья того или иного скота, посевных площадей под той или иной сельскохозяйственной культурой и т. п. Но когда речь идет об уровне рождаемости или смертности (в расчете на 1000 человек), об уровне безработицы, о средней урожайности зерновых или иных культур или других средних или относительных показателей, то применительно к ним использование термина «распределение по регионам» некорректно, так же как некорректно, на наш взгляд, строить для них кривую Лоренца и рассчитывать коэффициент Джини.

В свою очередь, по любым статистическим показателям, представленным в региональном разрезе, можно осуществить группировку регионов и построить вариационные ряды и гистограммы, характеризующие распределение регионов по выделенным группам с разным значением исследуемого показателя.

4. Отвечая на критические замечания оппонента, хотим сказать, что и у нас есть свои критические замечания в адрес статьи (2). Прежде всего это касается статистических примеров (и, в частности, графиков), приведенных в статье (2) с целью доказать правомерность использования кривой Лоренца и коэффициента Джини в оценке межрегиональных различий по любым показателям. Эти примеры не только не убедили нас в этом, а наоборот, заставили усомниться в самой постановке вопроса о возможности использования кривой Лоренца и коэффициента Джини в оценке межрегиональных различий.

Приведенные примеры трудно назвать таковыми, поскольку мы не видим ни исходных данных, ни расчетов показателей, по которым построены кривые Лоренца и определены коэффициенты Джини. В статье (2) приведены только конечные результаты и интерпретация их автором. Причем графики (рис. 1 и рис. 2) не лишены погрешностей. Так, на осях квадратов (где приведены шкалы значений показателей, по которым строится кривая Лоренца) следовало бы указать, что это не просто доли, а кумулятивные доли (регионов и исследуемых показателей).

Несколько удивляют названия рисунков: рис. 1 «Кривые Лоренца для распределения

регионов по уровню рождаемости, числу родившихся и численности населения»; рис. 2 «Кривые Лоренца для распределения регионов по душевому доходу, совокупному доходу жителей региона и численности населения» (2, с. 76). В такой редакции эти названия могут быть неправильно истолкованы, так как создают ложное представление, что кривые Лоренца характеризуют распределения регионов по всем перечисленным выше показателям. Не можем мы согласиться и с авторской интерпретацией представленных на рис. 1 и рис. 2 кривых Лоренца и соответствующих им коэффициентов Джини (по всем показателям) как показателей, отражающих межрегиональные неравенства (различия). Если эти кривые Лоренца построены по кумулятивным долям регионов и соответствующих суммарных показателей, то тогда они отражают неравномерность распределения, то есть концентрацию перечисленных показателей по регионам, а не межрегиональные различия (неравенства) в значениях показателя по регионам, как это постоянно трактуется в тексте статьи (2). При этом, согласно нашей позиции, кривые Лоренца, приведенные для таких показателей, как уровень рождаемости и среднедушевой доход, вообще не отражают ни концентрации, ни межрегиональных различий (то же самое касается и коэффициента Джини для этих показателей).

5. Наш оппонент очень любит термин «неравенство» и оценку его уровня с помощью коэффициента Джини. В конце статьи (2), желая удивить нас, автор пишет: «В заключение приведем экзотическую, но тем не менее вполне правомерную оценку межрегионального неравенства - «неравенство по неравенству», которая, с точки зрения положений рассматриваемой статьи², должна представляться совершенно дикой. Зададимся таким вопросом: какова степень различия российских регионов по величине внутрирегионального неравенства по доходам? Гистограмма соответствующего распределения приведена на рис. 4» (2, с. 78).

Не воспроизводя эту гистограмму, укажем, что на оси абсцисс в ней приведены интервалы значений коэффициента Джини (диапа-

зон коэффициента Джини), а на оси ординат - относительная частота. И озаглавлена гистограмма как «Распределение внутрирегионального неравенства по доходам». Никаких пояснений к гистограмме нет. Очевидно, автор считает, что сама гистограмма дает ответ на поставленный выше вопрос о степени различия российских регионов по величине внутрирегионального неравенства по доходам. Однако автору гистограммы мало; оказывается, нужна еще сводная оценка этих различий. Сразу после гистограммы читаем: «Сводную оценку степени различия регионов по внутрирегиональному неравенству дает коэффициент Джини, рассчитанный по коэффициентам Джини, которые характеризуют неравенство по доходам в регионах. Она оказывается низкой: 0,031. Таким образом, хотя неравенство по доходам внутри регионов довольно велико (в диапазоне от 0,359 до 0,481), между собой регионы в этом отношении в целом разнятся весьма мало» (2, с. 79).

Что можно сказать по поводу приведенного примера? Сама гистограмма, вопреки ожиданиям автора, вовсе не показалась нам «совершенно дикой». Вариационные ряды и отображающие их гистограммы могут строиться по любым признакам. Но название ее, мягко говоря, удивляет. Если мы правильно поняли, приведенная гистограмма отражает распределение коэффициентов Джини, рассчитанных для отдельных регионов по их значению (величине), или, что то же самое, распределение регионов по величине коэффициента Джини, характеризующего, по словам автора, неравенство по доходам в регионах (но не распределение неравенств).

Не оспаривая возможности построения такой гистограммы, хотим отметить, что, на наш взгляд, различия в размерах среднедушевого дохода по регионам более наглядно отразила бы просто гистограмма распределения регионов по величине среднедушевого дохода в них (то есть без расчета коэффициентов Джини по регионам).

Что касается вышеупомянутого показателя сводной оценки степени различия регионов по внутрирегиональному неравенству (в виде коэффициента Джини, рассчитанного по коэффициентам Джини, которые характе-

² Имеется в виду наша статья в журнале «Вопросы статистики» № 9 за 2015 г. (сноска наша: Г.Г., И.М.).

ризуют неравенство по доходам в регионах), то мы воздержимся от высказываний своего мнения по нему, дабы не давать повод дальнейшей дискуссии по данной теме.

6. При чтении статьи (2) невольно создается впечатление, что содержательная сторона показателей, используемых, в частности, в межрегиональных исследованиях, мало волнует автора. Для него коэффициент Джини – один из измерителей неравенств, а что он измеряет, неважно. И «хвала» коэффициенту Джини как измерителю неравенства поется и там, где он и не присутствует вовсе. Это можно проиллюстрировать следующим. Так, чтобы доказать нам, что коэффициент Джини давно используется в анализе межрегиональных неравенств применительно к среднему душевому и другим средним и относительным показателям, автор в качестве примера ссылается на статью 50-летней давности, в которой, по его словам, «оценивается межрегиональное неравенство в разных странах

по данным о доходах на душу населения регионов (правда, с применением коэффициента *вариации*, а не коэффициента Джини, но это значения не имеет)» (2, с. 72). (выделено нами: Г.Г. и И.М.).

Как говорится, комментарии излишни. Для нашего оппонента не имеет значения, что в упомянутом им исследовании межрегиональные различия по среднему душевому доходу оценивались с помощью коэффициента вариации, а не коэффициента Джини.

В заключение хотим отметить, что критические замечания К.П. Глущенко в адрес нашей статьи и его собственные высказывания по затронутым вопросам не изменили наших убеждений, изложенных в статье (1), и мы по-прежнему считаем, что кривая Лоренца и коэффициент Джини как индекс концентрации (характеризующий неравномерность распределения) могут использоваться для объемных (суммарных показателей), а не для средних и относительных величин.

ONCE MORE ON THE GINI COEFFICIENT AS AN INDICATOR OF THE CONCENTRATION

Galina L. Gromyko

Author affiliation: Lomonosov Moscow State University (Moscow, Russia). E-mail: gromyko@econ.msu.ru.

Irina N. Matyukhina

Author affiliation: Lomonosov Moscow State University (Moscow, Russia). E-mail: iri1256@yandex.ru

This article is a continuation of the discussion that started with the publication in «Voprosy statistiki» of the article «On the use of the Gini index in economic and statistical studies» by G.L. Gromyko, I.N. Matyukhina (issue №9 for 2015) and continued in the issue No. 2 for 2016, with the article by K.P. Gluschenko «On the issue of application of the Gini coefficient and other inequality indices». Both papers expressed opposing views on the application area for the widely recognized Gini coefficient.

In the present article the authors include response to critical feedback from the authors of the second article, along with criticism of their main key points. The difference in authors' approaches manifest in the following issues: application area for the Gini coefficient as a concentration index, directly associated with the Lorenz curve and intended to evaluate the uneven distribution of the studied index by groups; interpretation of «distribution» and «concentration» concepts in statistics; selection of a grouping variable when constructing variational series of distribution; meaningful interpretations of concentration and differentiation (difference) indicators and etc.

Keywords: Lorenz curve, Gini coefficient, concentration index, distribution of population by per capita income, distribution and grouping in statistics, average and relative values, differentiation indicators.

JEL: C10, C1.