

Реферат выполнен Евгением Андреевым

Краткое содержание статьи

## **Gini coefficient as a life table function: computation from discrete data, decomposition of differences and empirical examples.**

by Vladimir M. Shkolnikov, Evgeny M. Andreev and Alexander Z. Begun  
Demographic Research, 2003. 8:11, pp. 305-358

### **Коэффициент Джини как функция таблиц смертности: расчет на основе дискретных данных, декомпозиция различий и эмпирические примеры**

В настоящее время ожидаемая продолжительность жизни во многих странах весьма высока, и интересно узнать в какой степени это достижение равнодоступно для всех людей. По этой причине меры неоднородности популяции по продолжительности жизни привлекают растущее внимание. Коэффициент Джини – наиболее распространенный статистический индекс неоднородности или неравенства в общественных науках. Он широко используется как мера экономического неравенства между индивидуумами или домохозяйствами. Точно также коэффициент Джини может быть использован в качестве меры неравенства по продолжительности жизни, т.е. меры неоднородности популяции по возрасту в смерти.

Коэффициент Джини уже применялся в исследованиях как для измерения масштаба различий в средней продолжительности жизни между социальными группами, так и для измерения межиндивидуальных различий в возрасте смерти. Данная статья предлагает набор инструментов для измерения и анализа межиндивидуального неравенства в продолжительности жизни с помощью коэффициента Джини, который рассматривается как еще одна, дополнительная функция таблицы смертности.

Существует много подходов к расчету коэффициента Джини. Наиболее популярный среди экономистов метод основан на использовании кривой Лоренца. Кендалл и Стюарт [Kendall, Stuart, 1966] дали самое простое для понимания определение, согласно которому Джини коэффициент равен отношению средней межиндивидуальной разницы в возрасте смерти к величине средней ожидаемой продолжительности жизни. В терминах полной таблицы смертности, это определение соответствует формуле:

$$G_0 = \frac{1}{2(l_0)^2 e_0} \sum_{x=0}^{\omega} \sum_{y=0}^{\omega} d_x \cdot d_y \cdot |\bar{x} - \bar{y}|,$$

где  $x, y$  - возраст смерти,  $\omega$  - максимальная длительность жизни в популяции,  $d_x, d_y$  - числа умирающих в условном поколении таблицы смертности в возрастном интервале  $[x, x+1)$  и  $[y, y+1)$ , соответственно, из начальной совокупности новорожденных  $l_0$  (корня таблицы),  $\bar{x}, \bar{y}$  - средние возраста смерти в этих интервалах и  $e_0$  - ожидаемая продолжительность жизни при рождении.  $G_0$  – безразмерный показатель относительного неравенства. Если убрать деление на  $e_0$ , то получится показатель абсолютного неравенства, которые измеряется в годах жизни и может быть назван «средняя межиндивидуальная разница по продолжительности жизни» *AID*.

Однако с точки зрения простоты вычислений и использования тех же подходов, что и при расчете других показателей таблиц смертности, предпочтение может быть отдано формуле Ханада [Hanada, 1983]:

$$G_0 = 1 - \frac{1}{e(0) \cdot [l(0)]^2} \cdot \int_0^{\infty} [l(x)]^2 dx,$$

где  $l(x)$  – функция дожития. Эта формула очевидно аналогична основной формуле для расчета ожидаемой продолжительности жизни при рождении:

$$e_0 = \frac{1}{l(0)} \cdot \int_0^{\infty} l(x) dx.$$

Коэффициент Джини далеко не единственная возможная мера неравенства. Среди возможных индексов неравенства отметим такие, как расстояние между первым и третьим квартилем распределения по возрасту смерти  $IQR$ , дисперсия  $VAR$  и стандартное отклонение  $STD$  этого распределения, дисперсия логарифма длительности жизни  $VarLog$ , индекс энтропии распределения (индекс Тейла)  $T$ . Поэтому были выработаны критерии позволяющие оценивать предпочтительность того или иного индекса.

Так Ананд [Anand, 1983] сформулировал следующие основные свойства, которые желательны для любого индекса неравенства по уровню дохода: (а) независимость от численности, индекс не изменится, если общее количество людей изменяется без изменения в пропорциях личных доходов; (б) независимость от шкалы, индекс не изменяется, если все доходы изменяются пропорционально на одну и ту же величину; (в) условие Pigou-Dalton: любая передача части дохода от более богатого более бедному индивиду, которая не меняет их местами по уровню богатства, уменьшает величину индекса неравенства. Среди 6 рассмотренных в статье индексов неравенства, которые могли бы быть применены для измерения неравенства по продолжительности жизни, лишь два: коэффициент Джини и индекс энтропии  $T$  – удовлетворяют все трем условиям, но в отличие от индекса энтропии, восходящего к теории информации,  $G_0$  имеет достаточно простую интерпретацию. В случае смертности мужчин в США между 1950 и 1997 гг. тренды всех 6 рассмотренных индексов в основном совпадают, но в примере для мужчин России (1959-2000) единообразие динамики нарушается. Динамика индексов, повышенно чувствительных к уровню младенческой смертности, которая почти непрерывно снижалась в России, отлична от динамики тех, кто сильнее реагирует на смертность взрослых в динамике которой чередовались периоды снижения и роста.

Далее в статье детально рассматривается алгоритм расчета  $G_0$ , с учетом выявленной аналогии с  $e_0$ , в случае полных и кратких таблиц смертности. При расчете  $e_0$  по дискретным данным главный вопрос связан с определением для умершего в некотором интервале возраста  $[x, x + n)$  числа лет, прожитых им в этом интервале. Эта длительность характеризуется с помощью величины  $A_x$  - средняя доля возрастного интервала, прожитая людьми, которые умирают в пределах данного интервала. Если ряд  $A_x$  определен, то интеграл для расчета  $e_0$  заменяется суммой, которая в случае полной таблицы выглядит следующим образом  $e_0 = \sum_x [l_{x+1} + A_x(l_x - l_{x+1})]$ . Для расчета интеграла от квадрата функции

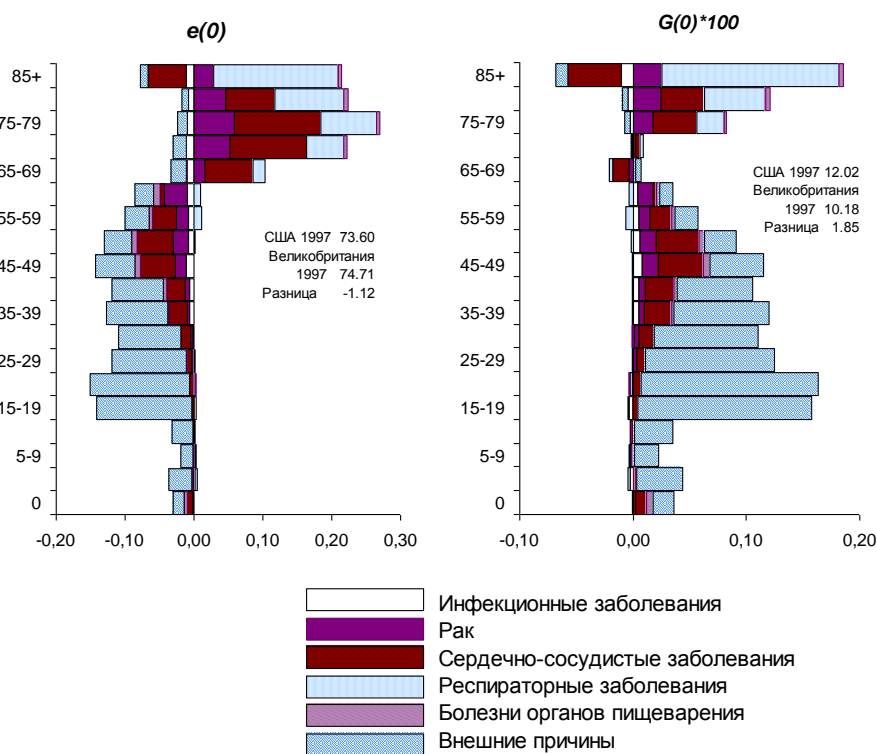
дожития  $[l(x)]^2$  вводится величина  $\hat{A}_x$ , которая близка, но не тождественна  $A_x$ . Значимые различия между этими двумя величинами существуют для возраста до 1 года, в случае кратких таблиц смертности, для открытого последнего возрастного интервала 85+. Предложенные приближенные расчетные формулы были успешно протестированы на основе 334 таблиц смертности для Японии, Франции, Швеции и США, почерпнутых из Berkeley Mortality Database.

Развивая дальше аналогию между методами расчета  $e_0$  и  $G_0$ , удастся найти формулы декомпозиции различия двух коэффициентов Джини по возрастам или по причинам смерти. Использован дискретный метод декомпозиции, получивший название

метода "пошаговой замены" ("stepwise replacement")<sup>1</sup>. Первоначально метод был применен к разложению различий в продолжительности жизни при рождении, где привел к известной формуле декомпозиции. Затем этот же алгоритм был применен к декомпозиции различий двух коэффициентов Джини, где дал принципиально новые формулы.

Рассмотрены примеры использования методов декомпозиции, которые не только иллюстрируют применение полученных формул, но и несут новую информацию. Так в США в 1900-1995 годах продолжительность жизни росла, а коэффициент Джини снижался. Изменение смертности во всех возрастных группах благоприятно влияло на рост продолжительности жизни, но особенно сильным было влияние смертности в возрастах 0-14 лет, снижение которой обусловило около 55% роста продолжительности жизни и мужчин и женщин. Одновременно это снижение обусловило более 70% снижения коэффициента Джини. Снижение смертности в старшей возрастной группе 65+ увеличило продолжительность жизни на 2,4 лет (более 9% общего роста), но привело к увеличению неравенства на величину равную примерно 5% изменения коэффициента Джини.

В качестве примера применения декомпозиции по возрасту и причинам смерти, проведено сравнение смертности мужчин в США и Великобритании в 1997 г. Продолжительности жизни мужчин очень близки - различие немного более 1 года в пользу Великобритании (или 1.4 % от  $e(0)$ ). Однако, коэффициента Джини в США был на 16%, чем Великобритании, что говорит об имеющемся в США существенном неравенстве по длительности жизни<sup>2</sup>. Рисунок 1 показывает компоненты различий между США и Великобританией.



**Рисунок 1. Декомпозиция различий в ожидаемой продолжительности жизни и в Коэффициенте Джини между Великобританией и США к возрасту и причине смерти: мужчины, 1997.**

<sup>1</sup> Подробно метод пошаговой замены изложен в статье [Andreev, Shkolnikov, Begun, 2002] также представленной в данной серии рефератов.

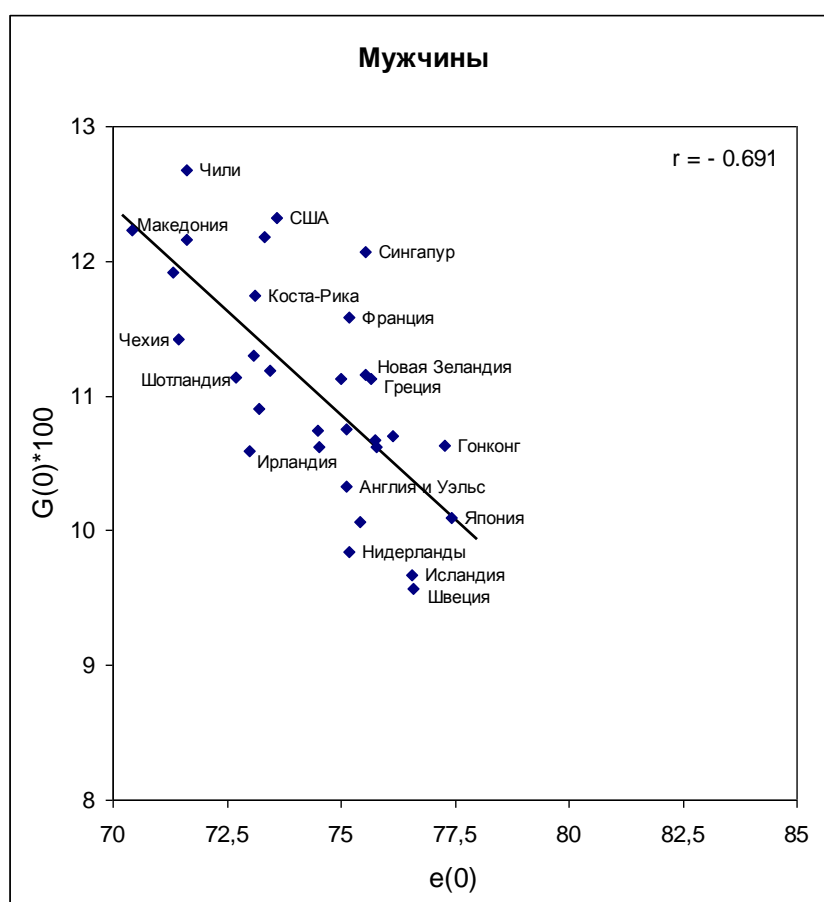
<sup>2</sup> В 2011 г. была опубликована статья, специально посвященная неоднородности смертности в США: Shkolnikov V. M. , Andreev E. M. , Zhang Z., Oeppen J., Vaupel, J. W. Losses of expected lifetime in the United States and other developed countries: methods and empirical analyses Demography, 48:1, pp. 211-239, которая также представлена в данной серии рефератов.

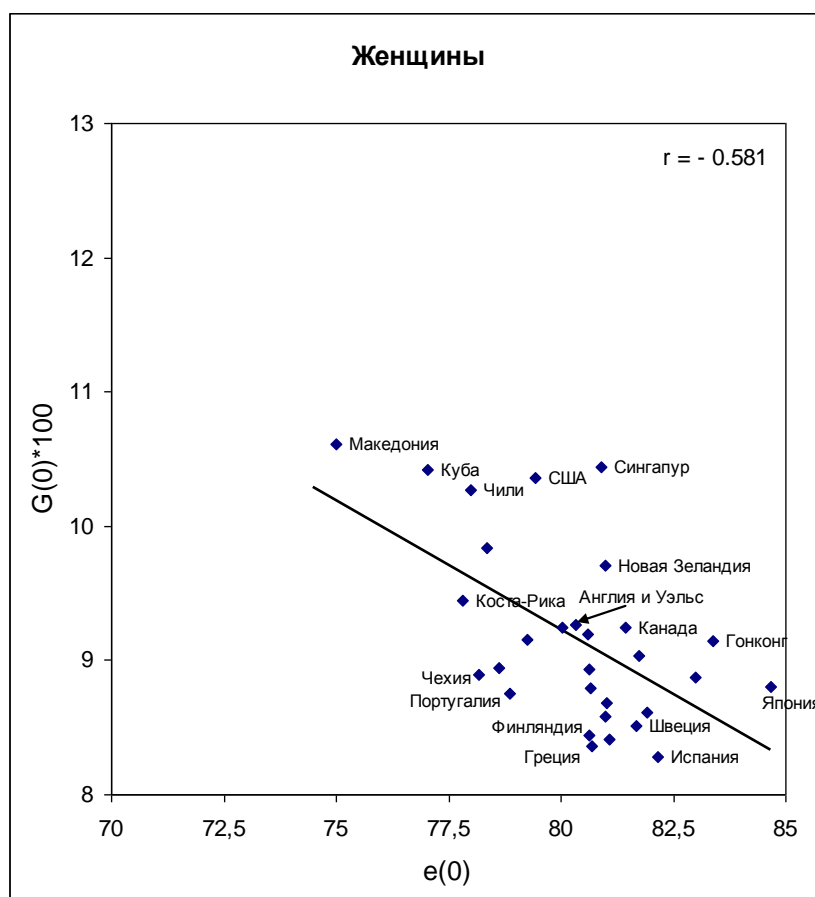
Преимущество Великобритании в продолжительности жизни мужчин связано главным образом с внешними причинами смерти в возрастах от 15 до 50 лет и несколько более низкой смертностью от болезней системы кровообращения и рака в возрастах от 40 до 59 лет. Однако, это преимущество почти уравновешено более низкой смертностью в США в возрастах 65 лет и старше от болезней системы кровообращения, респираторных заболеваний и рака. Вместе с тем, с точки зрения различий в коэффициенте Джини, роль повышенной смертности от внешних причин смерти в молодых возрастах существенно выше. Кроме того, и более низкая смертность в старости увеличивает уровень коэффициента Джини в США по сравнению с Великобританией.

В статье также рассмотрена возможность оценки вклада в различия между двумя величинами  $e(0)$  и между двумя величинами  $G(0)$  в зависимости не только от возраста, но и от социальной структуры населения.

Заключительная часть статьи посвящена анализу изменениям продолжительности жизни и коэффициента Джини во времени и пространстве.

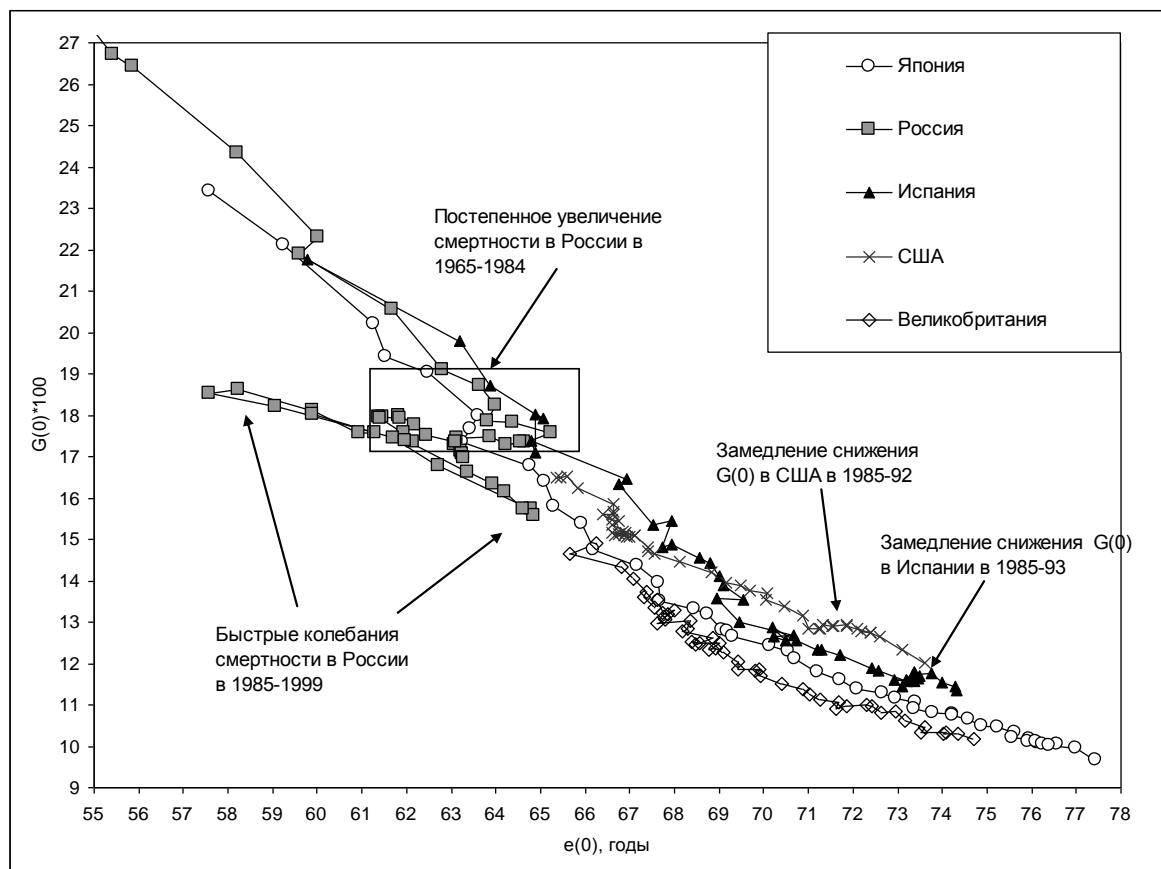
На рисунке 2 представлено соотношение между продолжительностью жизни для мужчин и женщин и коэффициентом Джини в 1996-99 гг. в 31 стране с продолжительностью жизни мужчин 70 лет или выше. Отчетливо видно, что с ростом продолжительности жизни коэффициент Джини убывает. Коэффициент корреляции для мужчин равен  $-0,69$  и для женщин  $-0,58$ .





**Рисунок 2. Соотношение между продолжительностью жизни и коэффициентом Джини для мужчин и женщин в 1996-99 гг. для 31 страны с продолжительностью жизни мужчин 70 лет и выше**

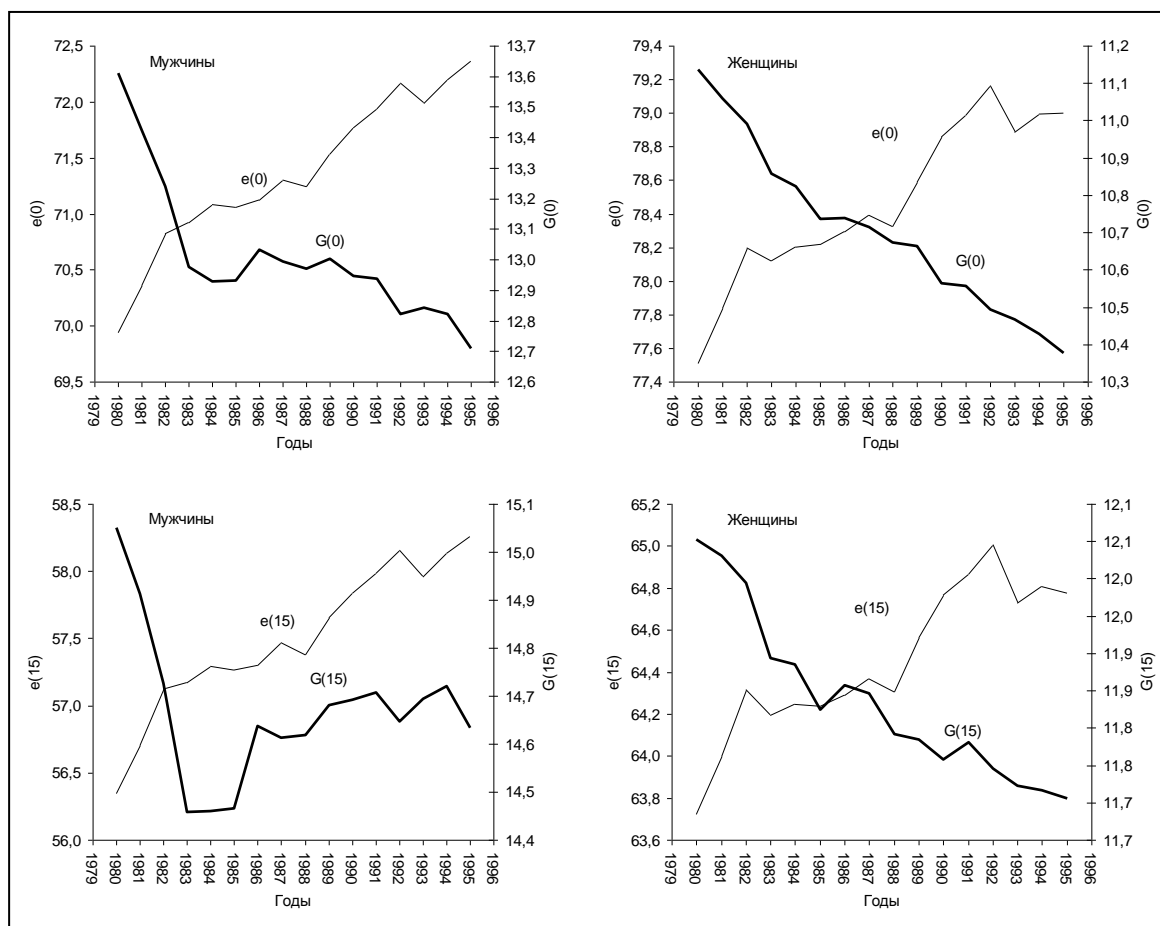
Траектории в координатах  $e(0)$  и  $G(0) \cdot 100$  для мужчин Великобритании, Испании, России, США и Японии в 1950-99 гг. представлены на рисунке 3. Траектории всех стран, кроме России, отражают устойчивый рост продолжительности жизни и уменьшение неравенства в длительности жизни. В 1950-е годы Великобритания и США опережали другие страны, но в конце периода лидером становится Япония. В конце 1980-х – начале 1990-х в Испании и США отмечалось прекращение снижения коэффициента Джини, но позже оно восстановилось. Россия выпадает из общего ряда. После впечатляющих успехов 1950-х, с середины 1960-х смертность мужчин увеличивалась. Последующие хаотические колебания соответствуют быстрому росту продолжительности жизни в годы антиалкогольной кампании и ее стремительному падению в первой половине 1990-х.



**Рисунок 3. Траектории в координатах  $e(0)$  и  $G(0) \cdot 100$  для мужчин Японии, России, Испании, США, и Великобритании в 1950-99 гг.**

Рисунок 4 представляет крупным планом динамику продолжительности жизни и коэффициента Джини мужчин и женщин в США в 1980-1996 гг. Ожидаемая продолжительность жизни мужчин увеличивалась в этот период непрерывно, в то время как коэффициент Джини не снижался между 1983 и 1990 годом. Если исключить влияние младенческой смертности и рассмотреть показатели для возраста 15 лет, то ситуация становится яснее: продолжительность жизни в 15 лет увеличивается, в то время как коэффициент Джини между 1983 и 1994 г. также увеличивается. В отличие от мужчин, для женщин ситуация кажется нормальной: и для новорожденных и в возрасте 15 лет продолжительность жизни увеличивается и коэффициент Джини снижается.

Применив метод декомпозиции можно доказать, что рост коэффициента Джини связан с ростом смертности мужчин в возрастах от 15 до 44 лет, главным образом из-за СПИДа и внешних причин в возрастах 15-29 лет. На рост коэффициента повлияло также уменьшение в смертности в самых старших возрастах.



**Рисунок 4. Тенденции в продолжительности жизни и коэффициента Джини в возрастах 0 и 15 лет для мужчин и женщин в США в 1980-95 гг.**

Главные выводы статьи – по своим математическим свойствам коэффициент Джини лучше других известных показателей соответствует базовым требованиям, предъявляемым к мерам неравенства. Динамика этих показателей в США и России с 1950-х показывает, что выбор меры неравенства может влиять на принципиальную оценку изменений. Коэффициент Джини, в отличие от некоторых других мер, не чрезмерно чувствителен к изменениям в детской смертности и соразмерно реагирует на изменения в смертности взрослых. Огромный массив вычислений свидетельствует об эффективности предложенного инструментария для расчета коэффициента Джини на основе показателей таблиц смертности и для декомпозиции различий двух коэффициентов по возрастам и причинам смерти.

Эмпирические примеры показывают, что, не смотря на тесную связь коэффициента Джини с продолжительностью жизни, некоторые изменения в смертности (например, рост смертности в молодых возрастах или снижение смертности в самых старших) могут разнонаправленно влиять на эти показатели.

### References

Anand, S. 1983. *Inequality and Poverty in Malaysia: Measurement and Decomposition*. New York: Oxford University Press.

Andreev, E.M., Shkolnikov, V.M. and Begun, A.Z. 2002. Algorithm for decomposition of differences between aggregate demographic measures and its application to life expectancies, healthy life expectancies, parity-progression ratios and total fertility rates. *Demographic Research*, 7:14, pp. 499-522

Kendall, M.G., Stuart, A. 1966. *The Advanced Theory of Statistics*. London: Charles Griffin.

Hanada, Kyo. 1983. A formula of Gini's concentration ratio and its application to life tables. *Journal of Japan Statistical Society*, 13: 95-98.