Наш семинар: математические сюжеты

Средние Джини

А. Б. Певный, С. М. Ситник

1. Широко известны три вида средних значений для пары положительных чисел a, b. Это *среднее арифметическое* $A(a, b)$, *среднее геометрическое* $G(a, b)$ и *среднее гармоническое* $H(a, b)$, которые определяются по формулам

$$A(a,b)=\frac{a+b}{2},\quad G(a,b)=\sqrt{ab},\quad H(a,b)=\frac{2ab}{a+b}.$$

Эти средние были введены ещё в Древней Греции в работах Пифагора и его школы, а также Никомаха [4, 15]. Источником введения средних для древнегреческих учёных стали пропорции, учение о которых активно развивалось, так как понятие пропорции использовалось тогда не только в математике, но также и в философии, скульптуре, архитектуре и живописи [4, 15]. Например, если найти величину x из нижеследующих пропорций, то как раз получим последовательно основные средние:

$$\begin{aligned}
\frac{a-x}{x-b}&=\frac{a}{a}=1 &&\Rightarrow\quad x=A(a,b),\\
\frac{a-x}{x-b}&=\frac{a}{x} &&\Rightarrow\quad x=G(a,b),\\
\frac{a-x}{x-b}&=\frac{a}{b} &&\Rightarrow\quad x=H(a,b).
\end{aligned}$$

Позднее были введены *среднее квадратичное*

$$Q(a,b)=\sqrt{\frac{a^2+b^2}{2}}$$

и общие *степенные средние порядка* r, которые для набора неотрицательных (при $r \geqslant 0$) или положительных (при $r < 0$) чисел определяются по формулам

$$\begin{gathered} M_r(a_1, a_2, \ldots, a_n) = \left(\frac{a_1^r + a_2^r + \ldots + a_n^r}{n}\right)^{1/r}, \\ M_0(a_1, a_2, \ldots, a_n) = \lim_{r \to 0} M_r(a_1, a_2, \ldots, a_n) = \sqrt[n]{a_1 \cdot a_2 \cdot \ldots \cdot a_n}. \end{gathered} \tag{1}$$

При этом справедливы формулы, выражающие основные средние через степенные. Например, для двух чисел

$$\begin{gathered} M_{-1}(a, b) = H(a, b), \quad M_0(a, b) = G(a, b), \\ M_1(a, b) = A(a, b), \quad M_2(a, b) = Q(a, b). \end{gathered}$$

Важнейшим свойством степенных средних является тот факт, что семейство этих величин монотонно возрастает по параметру r, или, как говорят, образует шкалу [1, 7, 10–12]. В частности,

$$H(a, b) \leqslant G(a, b) \leqslant A(a, b).$$

Эта самая известная из целого ряда теорем, в которой сравниваются между собой различные средние.

Наряду со степенными в различных задачах потребовалось использовать и другие типы средних [7, 11, 14]. В 1938 году итальянский математик и экономист Коррадо Джини (Corrado Gini, 1884–1965) ввёл новое семейство средних, которые впоследствии были названы его именем. Эти средние величины определяются так. Для любого набора положительных чисел $a = (a_1, a_2, \ldots, a_n)$ введём степенные суммы

$$S_p(a) = \sum_{i=1}^{n} a_i^p. \tag{2}$$

Тогда *средние Джини* $Gi_{p,q}(a)$, зависящие от двух вещественных параметров p и q, определяются по формулам

$$Gi_{p,q}(a) = Gi_{q,p}(a) = \begin{cases} \left(\dfrac{S_p(a)}{S_q(a)}\right)^{\frac{1}{p-q}}, & \text{если } p \neq q; \\ \exp\left(\dfrac{\sum_{i=1}^n a_i^p \ln a_i}{\sum_{i=1}^n a_i^p}\right), & \text{если } p = q \neq 0; \\ M_0(a), & \text{если } p = q = 0. \end{cases} \tag{3}$$

При этом нетрудно показать, что в соотношениях (3) второй случай получается из общего первого, а третий из второго при помощи предельных переходов. В частном случае, когда один из параметров равен нулю,

среднее Джини сводится к степенному среднему (1), а именно, справедливо соотношение
$$Gi_{p,0}(a_1, a_2, \dots, a_n) = M_p(a_1, a_2, \dots, a_n). \tag{4}$$
Мы далее рассматриваем общий случай, положив для определённости $p > q$.

Основным предметом изучения Коррадо Джини в экономической теории было имущественное неравенство, для его описания он ввёл знаменитый «*индекс Джини*», для вычисления и анализа которого использовались в том числе и различные классы средних. Это одно из основных понятий современной социальной статистики [8], а сам К. Джини — один из общепризнанных её создателей. Его результаты также известны в области статистического анализа проблем демографии. Джини использовал и пропагандировал результаты своего соотечественника Вильфредо Парето. Парето первым провёл конкретные расчёты, из которых следовало, что в его время всего 20 % населения владело 80 % национального богатства, и сделал вывод, что организованное таким образом государство ждёт быстрая и неизбежная гибель. Отметим, что сейчас в России реальные показатели индексов Джини ещё хуже (для неимущих). Развивая работы Парето, американский экономист М. О. Лоренц разработал теорию оценки разницы доходов населения по группам (кривая Лоренца), которая сейчас называется методологией Парето — Лоренца — Джини [8]. Отметим и третьего знаменитого итальянского математика и статистика XX века — Бруно де Финетти, разработавшего собственную оригинальную систему взглядов на основания теории вероятностей, автора известного в финансовой математике *парадокса де Финетти* (если капитал страховой компании конечен, то за бесконечное время она наверняка разоряется).

Несмотря на то, что по запросу в интернете на фамилии Парето, Джини или де Финетти будет выдано несколько тысяч ссылок, найти информацию о них самих практически невозможно. Это неумное замалчивание есть проявление так называемой политкорректности и связано с тем, что все они были сторонниками итальянского фашизма, в котором, вероятно, видели способ возрождения Италии. На идеях Парето воспитывался Муссолини, а Джини при нём занимал высокие посты и получал полную поддержку. Де Финетти прошёл путь от идейного сторонника фашизма до убеждённого коммуниста, читавшего лекции вьетнамским бойцам в джунглях под американскими бомбёжками. Разумеется, идеология фашизма во всех его проявлениях достойна однозначного осуждения, как и все её носители, кем бы сами они себя ни считали. Однако это не повод замалчивать значительный вклад этих учёных в общечеловеческую культуру и науку, тем более что они никак не запятнали себя непосредственным участием в преступлениях. Понятие оптимума по Парето, средних Джини и индекса

Джини, теория Колмогорова — Нагумо — де Финетти квазиарифметических средних, названная по именам её создателей, — всё это важнейшие и необходимые составляющие современной математики и статистики.

Важный частный случай средних Джини получается из (3) при $q = p - 1$. Эти средние были переоткрыты в 1971 г. Д. Лемером и имеют вид

$$Le_p(a) = \frac{\sum_{i=1}^n a_i^p}{\sum_{i=1}^n a_i^{p-1}}.$$

Их в советское время чуть позже Лемера также переоткрыл школьный учитель Ю. М. Фирсов [5].

2. Покажем, что средние Джини, так же как и степенные средние, образуют шкалу по своим параметрам, для них справедлива теорема сравнения. А именно, средние Джини увеличиваются, когда или параметр p, или параметр q, или оба эти параметра увеличиваются.

Теорема. *Пусть среди положительных чисел* $(a_1, a_2, \ldots, a_n)$ *есть хотя бы два неравных числа. Если выполнены условия* $p_1 > q_1$, $p_2 > q_2$, $p_2 \geqslant p_1$, $q_2 \geqslant q_1$ *и хотя бы одно из двух последних неравенств строгое, то справедливо неравенство*

$$Gi_{p_1,q_1}(a) < Gi_{p_2,q_2}(a). \tag{5}$$

Доказательство. Неравенство (5) в развёрнутом виде, записанное через степенные суммы (2), выглядит так:

$$\left(\frac{S_{p_1}(a)}{S_{q_1}(a)}\right)^{\frac{1}{p_1-q_1}} < \left(\frac{S_{p_2}(a)}{S_{q_2}(a)}\right)^{\frac{1}{p_2-q_2}}.$$

Оно упрощается, если от обеих частей взять логарифм:

$$\frac{1}{p_1 - q_1}(\ln S_{p_1}(a) - \ln S_{q_1}(a)) < \frac{1}{p_2 - q_2}(\ln S_{p_2}(a) - \ln S_{q_2}(a)). \tag{6}$$

Введём функцию $f(p) = \ln S_p(a)$. Тогда неравенство (6) переписывается так:

$$\frac{f(p_1) - f(q_1)}{p_1 - q_1} < \frac{f(p_2) - f(q_2)}{p_2 - q_2}. \tag{7}$$

Но (7) заведомо выполняется, если функция $f(p)$ строго выпукла. Действительно (см. рис. 1), левая часть (7) — это тангенс наклона секущей, проходящей через точки $(p_1, f(p_1))$, $(q_1, f(q_1))$, а правая часть — это тангенс наклона второй секущей. В условиях теоремы первый тангенс меньше второго тангенса (см. рис. 1).

Осталось установить строгую выпуклость функции $f(p)$, для чего достаточно установить неравенство $f''(p) > 0$, $-\infty < p < \infty$. Вычисление даёт

$$f''(p) = \frac{1}{S^2(p)}(S''(p)S(p) - (S'(p))^2).$$

Это выражение будет положительным, если выполнено условие

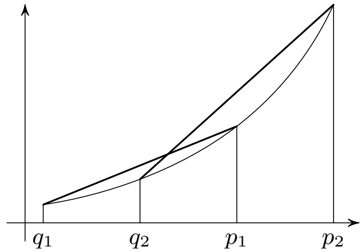

Рис. 1

$$(S'(p))^2 < S''(p)S(p). \tag{8}$$

В развёрнутом виде неравенство (8) эквивалентно следующему:

$$\left(\sum_{i=1}^{n} a_i^p \ln a_i\right)^2 < \sum_{i=1}^{n} a_i^p \sum_{i=1}^{n} a_i^p (\ln a_i)^2.$$

Но последнее соотношение следует из неравенства Коши — Буняковского для векторов с координатами $\sqrt{a_i^p}$ и $\sqrt{a_i^p}\ln a_i$.

Покажем, что действительно неравенство (8) всегда является строгим, хотя в использованном нами неравенстве Коши — Буняковского возможен и случай равенства. В этом случае должно существовать такое число λ, что

$$\sqrt{a_i^p}\ln a_i = \lambda\sqrt{a_i^p}, \quad i = 1, 2, \ldots, n,$$

а это возможно, только если все числа a_i равны, что противоречит условию теоремы. Следовательно, в (8) всегда будет строгое неравенство. Поэтому при сформулированных в условиях теоремы предположениях также строгими будут неравенства (7) и (5).

Это завершает доказательство теоремы. □

По мнению авторов, приведённое доказательство проще, чем в монографии Буллена [11, с. 249], в том числе и потому, что не содержит ссылок на результаты других работ.

В качестве следствия рассмотрим неравенства между средними Джини и более известными степенными средними, которые получаются с учётом соотношения (4).

Следствие. *Пусть среди положительных чисел $a_1, a_2, \ldots, a_n$ есть хотя бы два неравных. Тогда:*

1) *если $r \geqslant p > q$, $r > 0 > q$, то справедливо неравенство*

$$Gi_{p,q}(a_1, a_2, \ldots, a_n) < M_r(a_1, a_2, \ldots, a_n);$$

2) *если $p \geqslant r > 0$, $p > q > 0$, то справедливо неравенство*

$$M_r(a_1, a_2, \ldots, a_n) < Gi_{p,q}(a_1, a_2, \ldots, a_n).$$

В качестве упражнения предлагаем читателю самостоятельно доказать, используя определение (3), что выполнены соотношения

$$\lim_{p\to\infty} Gi_{p,q}(a) = \max_{1\leqslant i\leqslant n} a_i, \qquad \lim_{q\to-\infty} Gi_{p,q}(a) = \min_{1\leqslant i\leqslant n} a_i,$$

из которых следует, что среднее Джини (с учётом доказанной теоремы о монотонности по параметрам!) находится между минимальным и максимальным из чисел a_i:

$$\min_{1\leqslant i\leqslant n} a_i \leqslant Gi_{p,q}(a) \leqslant \max_{1\leqslant i\leqslant n} a_i,$$

и, таким образом, является настоящим средним.

3. Средние значения достаточно широко используются в химии, в частности в теории полимеризации. В качестве примеров укажем известные монографии [2, 3, 9]. Так, в монографии [2] в главе о молекулярно-массовом распределении изучаются различные способы усреднения по массе молекул и по числу частиц. Вводится и изучается весовое среднее арифметическое [2, с. 127], для средневязкостной молекулярной массы вводится весовое степенное среднее вида (1), см. [2, с. 130]. С математической точки зрения, при этом без доказательства используется монотонность весовых степенных средних по параметру (свойство шкалы средних, см. [1, 7, 10–12]), а также неравенство между весовыми средними и невесовым средним арифметическим. При изучении связей механизмов полимеризации с молекулярно-массовым распределением вводятся и изучаются свойства интегральных аналогов средних Джини $Gi_{3,2}, Gi_{2,1}$ [2, с. 184–185], с их использованием определяется так называемый коэффициент полидисперсности Шульца.

В монографии [3] при изучении физических методов исследования макромолекул в растворах вводится так называемое вискозиметрическое среднее, которое выражается через молекулярные веса макромолекул с помощью среднего Джини $Gi_{s+1,1}$ [3, с. 36].

Различные средние величины также широко используются в известной монографии [9]. Так, на с. 23–24 для средних значений молекулярных весов вводятся последовательно весовое интегральное среднее произвольного порядка, вес которого равен численной функции распределения; среднее Джини, совпадающее со средним Лемера $Gi_{q,q-1} = Le_q$, которое называется q-средним молекулярным весом, а также степенное среднее. На с. 67 при статистическом анализе механизмов фракционирования вводится среднее Джини — Лемера $Gi_{2,1} = Le_2$. На с. 83 при изучении гидродинамических средних весов и связанных с ними критериев полидисперсности вводятся понятия среднедиффузионного и среднеседиментационного весов, которые выражаются через весовые степенные интегральные средние некоторых отрицательных порядков, а также весовое интегральное среднее Джини $Gi_{1,1-b}$, где b — это некоторая характеристика, изменяющаяся в пределах $0 < b < 1$. При этом без строгого доказательства, с необычной для математика формулировкой «обычно это верно» утверждается, что указанное среднее Джини всегда больше среднего арифметического. Отметим, что строгое доказательство этого утверждения вытекает из приведённого выше следствия при выборе параметров $p = 1$, $q = 1 - b$, $0 < q < 1$, $r = 1$. На с. 83–84 вводится некоторая величина, связанная с молекулярными весами, которая выражается через среднее Джини — Лемера $Gi_{2-b,1-b} = Le_{2-b}$, а на с. 229 для минимальной величины эффективного параметра получено выражение через среднее Джини $Gi_{3/2,-3/2}$. Вопросы предельного поведения средних Джини в терминах отношения степенных средних рассматриваются также на с. 252 при изучении возможностей сведе́ния процессов полимеризации к основным статистическим классам распределений.

Таким образом, различные классы средних значений, включая средние Джини, находят важные применения в химии при рассмотрении многочисленных вопросов теории полимеризации молекул. При этом ряд используемых в химической литературе нестрогих соотношений между средними может быть строго обоснован с использованием доказанных в настоящей статье неравенств между средними Джини и степенными средними.

4. В заключение отметим, что так как в нашем доказательстве используется неравенство Коши — Буняковского, то можно получить некоторые усиления доказанных неравенств с использованием уточнений неравенства Коши — Буняковского, полученных в [7, 14]. Авторы также использовали неравенства для средних Джини при доказательстве интересного неравенства А. С. Гаспаряна между однородными степенными формами [6, 13].

Список литературы

[1] *Беккенбах Э., Беллман Р.* Неравенства. М.: Мир, 1965.

[2] *Берлин Ал. Ал., Вольфсон С. А., Ениколопян Н. С.* Кинетика полимеризационных процессов. М.: Химия, 1978.

[3] *Волькенштейн М. В.* Конфигурационная статистика полимерных цепей. М.-Л.: АН СССР, 1959.

[4] *Джини К.* Средние величины. М.: Статистика. 1970.

[5] *Калинин С. И.* Средние величины степенного типа. Неравенства Коши и Ки Фана. Киров: Изд. ВятГГУ, 2002.

[6] *Певный А. Б., Ситник С. М.* Об одном неравенстве А. С. Гаспаряна // Проблемы матем. анализа. 2014. Вып. 77. С. 159–162.

[7] *Ситник С. М.* Уточнения и обобщения классических неравенств // Итоги науки. ЮФО Серия «Математический форум: Исследования по математическому анализу», ред. Ю. Ф. Коробейник, А. Г. Кусраев. Владикавказ, 2009. Т. 3. С. 221–266.

[8] Социальная статистика / Ред. И. И. Елисеева. М.: Финансы и статистика, 2003.

[9] *Френкель С. Я.* Введение в статистическую теорию полимеризации. М.-Л.: Наука, 1965.

[10] *Харди Г. Г., Литтльвуд Дж. Е., Полиа Г.* Неравенства. М.: ИЛ, 1948.

[11] *Bullen P. S.* Handbook of means and their inequalities. Dordrecht: Kluwer Academic Publishers, 2003. (Mathematics and its Applications (East European Series); V. 560).

[12] *Mitrinović D. S., Pečarić J. E., Fink A. M.* Classical and new inequalities in analysis. Dordrecht: Kluwer Academic Publishers, 1993. (Mathematics and its Applications (East European Series); V. 61).

[13] *Pevnyi A. B., Sitnik S. M.* On Gasparyan's inequality // J. Math. Sci. 2015. V. 205, № 2. P. 304–307.

[14] *Sitnik S. M.* Generalized Young and Cauchy — Bunyakowsky inequalities with applications: a survey. http://arxiv.org/abs/1012.3864.

[15] *Toader Gh., Toader S.* Greek means and the arithmetic–geometric mean. RGMIA Monographs, Victoria University, 2005. http://rgmia.org/monographs/toader.htm

А. Б. Певный, Сыктывкарский государственный университет

pevnyi@syktsu.ru

С. М. Ситник, Воронежский институт МВД России

pochtasms@gmail.com