

タイトル	平均概念 : ジニ 『平均論』 (ミラノ, 1958年)断章
著者	木村, 和範
引用	季刊北海学園大学経済論集, 56(3): 135-150
発行日	2008-12-25

《研究ノート》

平均概念について

— ジニ『平均論』（ミラノ、1958年）断章 —

木 村 和 範

はじめに

1. ジニ『平均論』刊行の経緯と若干の基礎概念
 - (1) 刊行の経緯
 - (2) 基礎概念
 2. キズィーニの平均概念とジニの批判
 - (1) キズィーニの定義
 - (2) ジニの批判
- むすび

はじめに

数理統計学の教科書では、一般に、まず度数分布（図・表）が示され、それにかんする相対度数や累積相対度数が計算される。次に、メディアン（中央値）やモード（最頻値）などの「位置上の『平均』」がその度数分布の「代表値」として取り上げられる。その後の中心論点は「計算的『平均』」である。なかでも相加平均が重視されている。それは、散らばりの尺度として多用される分散が、平均偏差の平方にかんする相加平均（標準偏差はその平方根）であるからだけでなく、原系列とあてはめ線との平均的な乖離を最小とする誘導統計値としての回帰係数を算出するときにも相加平均を用いるために、数理統計学的手法の定番とも言うべき分散（標準偏差）や回帰係数を取り上げる前段で相加平均に言及しておく必要があることによる。

相関係数は2つの回帰係数の相乗平均としても定義される。このために、相加平均とならんで相乗平均が取り上げられることもある。

相乗平均は、平均経済成長率の算出に不可欠の概念であるが、その説明の丁寧さに欠ける教科書も少なくない。

相加平均や相乗平均と並んで、調和平均と言われる「計算的平均」もある。これは、実数の系列を構成する各項の値の逆数にかんする相加平均の逆数と定義されている。確率論基調の統計理論が主流になるに伴い、たとえ調和平均を取り上げている教科書でも、公式や例題、計算問題が掲載されるにとどまり、調和平均の数理的意味に言及されることは珍しくなった。

以上の記述的な統計的方法に次いで、数理統計学の教科書は推測的な統計的方法（推定や検定など）を取り上げている。そこでは、母集団と標本、母集団分布と標本分布にかんするさまざまな確率分布が解説の対象となる。その際、確率についての基礎知識が必要であることから、教科書では、記述的な統計的手法を取り上げることなく、ベン図を用いた確率の解説に始まり、確率の加法定理と乗法定理、ベイズの定理（逆確率）、中心極限定理などに紙幅が充てられ、その後に推測的な統計的方法（推定や検定）に論が進められることもある。推定や検定では、標本分布が数理の要となっているために、母集団—標本図式から叙述を始める教科書もある。いずれの教科書においても、相加平均は分散と並んで重要なパラメータであると規定され、推測すべき母集団特性値の代表格として重要な位置を

占めている。

このように、記述的と推測的とを問わず、統計学における平均と平均操作の位置づけを考えてみると、その割には、この国では平均の意義を考察した論文や書物の刊行が少ないことに気づく。そのようななかで、ジー・ジェック(岡崎文規訳)『統計的中数値論』(有斐閣, 1926年)、リャブーシキン(是永純弘訳)『統計学における平均』(モスクワ, 1954年)¹⁾、成島辰巳『社会科学のための平均論——学説史的研究——』(法政出版, 1995年)は異彩を放っている。

平均にかんしては、この他にも、1958年の刊行当初からこの国の統計学界が注目してきた著書がある。それはジニの『平均論』²⁾である。その全体を概観すれば、ジニが明言しているように、「数学の観点」(2)から平均を考察していることが分かる。その点で社会科学における平均(操作)の有効性を検討している上記三著作(ジー・ジェック、リャブーシキン、成島の著書)とは対照的である。

以下では、統計と統計理論における平均概念の意義に鑑みて、ジニ『平均論』を取り上げる。ただし、それは500頁に及ぶ大著であるために、ここでは、平均の概念規定をとり上げて、その一部に言及することしかできない。

1. ジニ『平均論』刊行の経緯と若干の基礎概念

(1) 刊行の経緯

1884年5月23日にトレヴィーゾ市

(Treviso)(ヴェネツィアより約30km北方)近郊のモッタ・ディ・リヴェンツァ(Motta di Livenza)で誕生したコッラド・ジニ(Corrado Gini)は、1965年3月13日早朝に、書斎の机で死亡しているのを発見された。その机上には書きかけの書簡があったという。80年あまりの生涯であった。23歳のときに最初の論文「小数法則」(1907年)³⁾を執筆し、半世紀を超える研究生活を通じて著書87冊(うち3冊はその印刷中に死亡したために生前には未公刊)を著し、論文は827篇に及んでいる⁴⁾。

標本調査の分野では、ジニは、代表標本の選出に失敗した「あのイタリアの統計学者」(J. ネイマン)とも言われている⁵⁾。しかしながら、彼は1950年代のイタリア統計学界では「イタリア学派」を率いた泰斗として知られ、ジニ係数の考案者として今日なおその令名を馳せている。この国では、田口時夫⁶⁾がジニの集中理論にかんする一般化を試みたが、その所説は、ジニ理論を拡充するものとして国際的な場で論議の対象となっている⁷⁾。

3) Gini, C., "La legge dei piccolo numeri," *Giornale degli Economisti*, Serie II, Vol. XXXV, 1907.

4) Castellano, Vittorio, "Corrado Gini: a Memoir," *Metron*, Vol. XXIV, N.1-4, p.3. ジニの文献目録は pp.36ff. に収録。

5) ネイマンのジニ批判が正当かどうかについては、木村和範『標本調査法の形成と展開』(北海道大学図書刊行会, 2001年)第6章参照。

6) 田口時夫『経済分析と多次元解析——新しい計量空間の形式と展望——』東洋経済新報社 1984年。なお、①同「ジニ統計学の数学的性質」『統計学』第65号, 1993年; ②吉田忠「多次元集中局面の統計学——田口時夫氏の多次元集中曲面による集団構造解析——」『統計学』第83号, 2002年も参照。

7) Arnold, Barry C., "The Lorenz curve: Evergreen after 100 years," in: Gianni Betti and Achille Lemmi (ed.), *Advances on Income Inequality and Concentration Measures*, London and New York 2008, Chap.2, p.21.

1) この訳書は法政大学出版局からタイプ印刷で発行されたが、刊行年は不詳である。

2) Gini, Corrado, *Le Medie*, in collaborazione con Gustavo Barbensi, Luigi Galvani, Stefania Gatti, Ernesto Pizzetti, Milano 1958. 以下、本文で()内に記したローマ数字と算用数字は、この著書の頁を示す。

さて、本稿で取り上げる著作『平均論』の序文はその刊行の前年(1957年)7月に執筆された。それによれば、第2次世界大戦以前に、数学者や統計学者によって平均にかんする著作が多数、執筆されていた。そのような学界の状況を反映して、1938年にボローニャ市で開催された第27回イタリア科学振興協会研究大会(XXVII Ruinione della Società per il Progresso delle Scienze)⁸⁾や1939年イタリア統計協会(Società Italiana di Statistica)ピサ研究大会において、平均にかんする研究報告がなされた(V-VI)。平均をめぐる論点は多岐に渡るため、それを整理し、体系化する必要があるとジニは考えた。

しかし、ジニの構想を実現すべく、平均にかんする組織的な研究が始まったのは戦後(1950-51年)のことである。当時、ローマ大学統計科学・人口論・保険数理学部(Ia Facoltà di Scienze Statistiche, Demografiche ed Attuariali)を統括する職にあったジニのもとに少数の有志が集まって、平均にかんする研究会が発足した。その後、この研究会に参集する研究者の数は次第に増えて、その研究成果が一書にまとめられることになった。それが、ジニ『平均論』(ミラノ、1958年)である。ジニが執筆した序文にはこの著作が上梓されるに至った経過と平均にかんする前述の研究会における多数の主要メンバーの名前が挙げられている。そのみならず、この書物のタイトル頁には、その研究会で主導的な役割を果たしたグスタヴォ・バルペンシ(Gustavo Barbensi)のほかに、ルイジ・ガルバーニ(Luigi Galvani)、ステファニア・ガッティ(Stefania Gatti)、エルネスト・

ピッツェッティ(Ernesto Pizzetti)の名前が、ジニの協力者として掲げられている。その名を挙げられた協力者の業績と本文のなかで引用・参照されている文献とを対応させると、これらの人々は単なる協力者ではなく、共著者であることを推測させる。また、印刷に回す直前の最終原稿の編集は、ローマ大学統計学講座(cattedra di Statistica dell'Università di Roma)で助手の職にあったガッティが担当した。それが、ジニの校訂を経て公刊された(VII)。このような経緯から、『平均論』はジニの単著であるというよりは、ジニの構想以来、「20年を超える長く、しかも忍耐を要した研究の成果」(VIII)であり、ジニ学派の1つの到達点を示す共同労作と見なすことができる。

(2) 基礎概念

① 計算的平均と位置上の平均

通常、われわれは「平均」を①相加平均や相乗平均のような「計算的平均」と②メディアンやモードのような「位置上の平均」に分類している(図1)。「位置上の平均」であっても、計算操作によって導出されることが多い。むしろ統計値集団が階級区分された度数分布としてあたえられるときには、メディアンやモードは計算式によらなければ、算出することができない(本稿末尾数学注参照)。そのために、「位置上の平均」と言えども、一種の「計算的平均」ではないのかという疑問が生ずる。

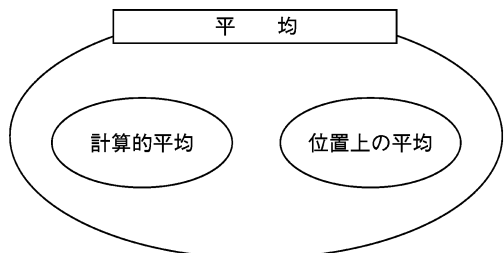


図1 平均の一般的な分類

8) Barbensi, Gustavo, "Cenni sulla trattazione monografica delle medie statistiche," *Atti delle XXVII Ruinione delle Società Italiana per il Progresso delle Scienze (Bologna, 4-11 settembre 1938)*; cf. Gini, C., "L'evoluzione del concetto di media," *Metron*, Vol.XVI, N.3-4, 1952, p.3.

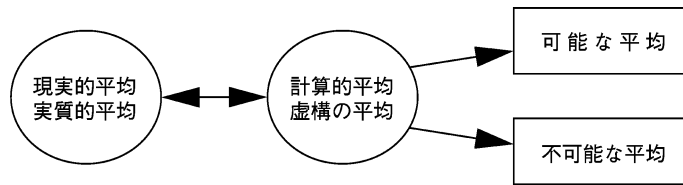


図2 ジニの分類

この疑問にたいするジニの見解は明快である。ジニによれば、実数の系列（統計系列を含む）があり、その系列について平均を算出する場合、実際に、その値が原系列を構成する項のなかに見いだしうるとき、その平均（「所与の分布を構成する項の1つに対応する平均」(444)）を「現実的平均ないし実質的平均 (medie reali o effettive)」と言う。要するに、「現実的平均 (実質的平均)」とは、原系列のなかにもとめられた平均値と同一の値をもつ項が存在する場合の平均である (63)。

これにたいして、算出された平均の値が原系列に存在しない場合、その平均を「計算的平均ないし虚構の平均 (medie di conto o fittizie)」と言う。たとえば7世帯の資産 (1, 12, 54, 130, 2600, 5800, 10000 [単位は千ドル]) について算出される相加平均は265万6700ドルであり、この値は「当の分布に帰属しない値によって表される平均」(444)である。このように、原系列のなかには見出し得ない「新しい値」が計算によって平均とされる場合に、それを「計算的平均 (虚構の平均)」と言う (63)。したがって、ジニによれば「計算的平均」の概念は「位置上の平均」と対照的に区別されることがない (図2)。

ジニは、さらにこの「計算的平均」を「可能な虚構の平均 (medie fittizie possibili)」と「不可能な虚構の平均 (medie fittizie impossibili)」とに分類した。前者の平均は、原系列のなかには存在しない数値ではあるが、可能性としてその存在に疑念の余地がない数値である。たとえば、先にもとめた平均値265万6700ドルは世帯の資産として把捉される

可能性がある。したがって、これは「可能な虚構の平均」である。

これにたいして、後者の平均（「不可能な虚構の平均」）としては、たとえば2人、3人、3人、4人、5人、5人、7人からなる7世帯の平均世帯員数 (4.14人) がある。ジニによれば、「1世帯が4.14人で構成されるということは絶対にあり得ない」からである (445)。

② 解析的平均と非解析的平均

ジニはメディアンやモードなどを「位置上の平均 (medie di posizione)」と言い、これを「非解析的平均 (medie non analitiche)」とも言っている (図3)。

これに対比される平均が「解析的平均 (medie analitiche)」である。この「解析的平均」には、通常、われわれは「計算的平均」という用語を当てている。本稿のはしがきでもそのような意味でこの言葉を使用した。

さて、ジニ『平均論』によれば、相加平均、相乗平均、調和平均などがこの「解析的平均」の具体例である。したがって、ジニの用語法とこの国における通常用語法との違いを要約すれば、次のようになる。

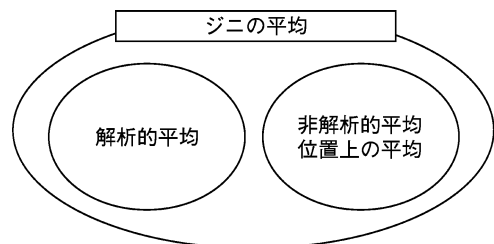


図3 2種類の平均

- i. 通常、われわれが用いる「計算的平均」はジニの「解析的平均」に該当する。ジニも「計算的平均」という概念を用いているが、それはわれわれの用法とは異なっている。
- ii. 「位置上の平均」については、ジニの用語法と通常用語法との間に齟齬はない。ただし、ジニは「位置上の平均」を「非解析的平均」とも言っている。

このような違いを確認した上で、以下ではさらにジニの用語法を検討する。彼によれば、「解析的平均」 M は、一定の制約を受けるが⁹⁾、基本的には系列 x_1, x_2, \dots, x_n の関数

$$M = f(x_1, x_2, \dots, x_n) \quad (1)$$

として表現することができる (64)。これにたいして、「非解析的平均」は(1)式のように表現されることなく、「[系列を構成する] 諸項の値にかんする数式表現が不可能であって、所与のすべての項からなる単調系列 (successione monotona) のなかに占める位置に依存しているので、位置上の平均とも言われている」(64。ただし、強調はジニによる)。

ジニによれば、「解析的平均」は「固定的 (fermo)」であるのにたいして (64)、「非解析的平均ないし位置上の平均」には「弛緩的 (lasco)」という特徴がある (102)。ここに「固定的」とは、系列ごとに平均の数値が定まっていて、1つの項の値が変化すれば、そのときの平均はもとの平均とは違った値になることを意味する。たとえば系列が 1, 2, 3, 4, 5 (系列 I) のときの相加平均は 3 であるが、末項の値が変化して 1, 2, 3, 4, 10 (系列 II) になるとき、相加平均は 4 になる。このようにジニの「固定的」という用語は、「平均」が項の値の変化にたいして「鋭敏」に反応する

という特性を表す。

次に、総度数を半分にするメディアン Me を取り上げて「位置上の平均」の特性とされる弛緩性について考える。上に掲げた系列 I と II のいずれにおいても Me は同一の 3 である。この場合に Me の値は変わらない。項の値の変化にたいして Me は不感的である¹⁰⁾。メディアンだけでなく、モードを含めたさまざまな「位置上の平均」には「概して」このような特性が見られる。このように「位置上の平均」は項の変化にたいして鋭敏さに欠けることから、ジニはその特性を「弛緩的」と表現した。

ただし、「固定的」と「弛緩的」という平均の特質にかんするジニの叙述は必ずしも明快ではない。「解析的平均」に典型的とされる特性 (固定性) が、一部の「非解析的平均」にも見られ、また逆に「非解析的平均」の特性とされる弛緩性が一部の「解析的平均」にも見られるという趣旨の指摘がある¹¹⁾。ここでは、それらの特性が、それぞれの平均に「概して」見られるにすぎないことに留意したい (102)。

2. キズィーニの平均概念とジニの批判

(1) キズィーニの定義

ジニ『平均論』では、同書の執筆者 (集団) と異なる意見をもつオスカル・キズィーニ (Oscar Chisini) を唯一の例外として、その名を挙げ批判している。この特別な取り扱

9) 「一定の制約」という文言に傍点を付した含意については次節で述べる。

10) メディアンが項の値の変化にたいしてつねに不感的という訳ではない。項の変化が「平均」に反映されることもある。たとえば、系列 I の各項が +1 ずつ増加するとき、メディアンは 3 から 4 に変化する。

11) たとえば、系列 I において両端項の 1 と 5 がそれぞれ 2 と 4 に変化するとき、相加平均は、項の値の変化にかかわらず、同一の 3 となる。

いは、彼の平均概念が、その根幹においてジニの定義とは異なっているためである。以下、キズィーニの論文「平均の概念について」(1929年)¹²⁾にもとづいて、その見解を紹介する。

キズィーニによれば、当時、平均概念についてはコーシー (Augustin Louis Cauchy [1789-1857]) による定義¹³⁾が一般的に採用されていた。コーシーは、「諸量の平均とは、それらの諸量における小さい量と大きい量との間にある新しい量のことである」¹⁴⁾と規定した。すなわち、系列を構成する諸項の量的規定性のなかで、もっとも小さな量的規定性よりも小さくなく、もっとも大きな量的規定性よりも大きくない、何らかの大きさの量的規定性をもって平均と見なすとコーシーは考えた。これが、ジニ『平均論』においても基本的には踏襲されている。

これにたいして、キズィーニは、コーシーの定義が数学的な厳密性に欠けると批判した。そして、平均には「解析的な表現」が必要であると考へて、次のように述べた¹⁵⁾。

「同質の量 (*grandezza omogenee*) を表す任意の n 個の従属変数 x_1, x_2, \dots, x_n にかんする関数

$$y = f(x_1, x_2, \dots, x_n) \quad (2)$$

があたえられるとき、関数 f にかんする x_1, x_2, \dots, x_n の平均とは、 x_1, x_2, \dots, x_n の代わりになって、その関数とまったく

同一の値をあたえる数 M 、すなわち

$$f(M, M, \dots, M) = f(x_1, x_2, \dots, x_n) \quad (3)$$

となるような数 M のことを言う」(強調はキズィーニによるが、式番号は引用者による)。

キズィーニは、相加平均、相乗平均、平方平均を例にして、(3)式の妥当性を主張した。(3)式の M は、一見すれば、ジニの「解析的平均」と同一であるかに見える。しかし、そうではない。先に「解析的平均」を表現するために

$$M = f(x_1, x_2, \dots, x_n) \quad (1)[再掲]$$

を掲げたとき、ジニの「解析的平均」 M は、「一定の制約」のもとで、かかる関数表現が可能であると指摘した。この「一定の制約」という文言によって、(1)式が平均の定義式となりうるには、それなりの条件を満たしていなければならないと考えたジニの主旨を表現したつもりである。そして、これによって「解析的平均」が一般にキズィーニの(3)式で定義されることはないことを示そうとした。

それでは、この「一定の制約」とは何か。また、(3)式が平均を一般的に規定する定義にはなり得ないとジニは考へ、キズィーニを批判したのは何故か。このことについては、項を改めて述べる。

(2) ジニの批判

ジニが、キズィーニの定義を受容できないと考えた第1の理由は、(3)式が「非解析的平均(位置上の平均)」を包含しないということであろう。メディアンやモードなどの「位置上の平均」は、原系列を従属変数とする関数として表現できないからである。

第2に、たとえ「解析的平均」であろうとも、キズィーニの定義式にもとづけば、原系

12) Chisini, Oscar, "Sul concetto di media," *Periodico di Matematiche*, Serie IV, Volume IX, N. 2, 1° marzo 1929 [Chisini (1929)].

13) ①Cauchy, A.L., *Cours d'analyse de l'Ecole Royale Polytechnique*, Première partie: *Analyse algèbre*, Paris 1821; ②ditto, *Œuvres complètes*, II^e Série. T. III, Paris 1897.

14) ただし、引用は Chisini (1929), p.106 による。

15) Chisini (1929), p.108.

列を構成する項のなかの最大値よりも大きな値が平均としてあたえられうることを否定できない。この例としてジニは「反調和平均 (la media antiarmonica)¹⁶⁾」 M を取り上げた

(60)。2つの実数 a_1 と a_2 にかんする反調和平均 M は

$$M = \frac{a_1^2 + a_2^2}{a_1 + a_2} \quad (4)$$

16) ピュタゴラス学派 (古代ギリシャ) は、 $a > b > c > 0$ または $0 < a < b < c$ について、

$$\frac{a-b}{b-c} = \frac{a}{c} \quad (1)$$

を満たす比を「算術的比例 (proporzione aritmetica)」と名づけた。そして、

$$\frac{a-b}{b-c} = \frac{b}{c} \quad (2) \quad \text{または} \quad \frac{a-b}{b-c} = \frac{a}{b} \quad (2')$$

を満たす比を「幾何学的比例 (proporzione geometrica)」と言い、

$$\frac{a-b}{b-c} = \frac{a}{c} \quad (3)$$

を満たす比を「調和的比例 (proporzione armonica)」と命名した。今日では、以上の3つの比は「古典的比例 (proporzione classiche)」と言われている (Gini, C., *Le Medie*, Milano 1958, p.2)。

さらにピュタゴラス学派は

$$\frac{b-c}{a-b} = \frac{a}{c} \quad (4) \quad \text{または} \quad \frac{a-b}{b-c} = \frac{c}{a} \quad (4')$$

を満たす比を「反調和的比例 (proporzione antiarmonica)」あるいは「小反対的比例 (proporzione subcontraria)」と名づけた (*op. cit.*, p.13)。

ここで、①式を b について解けば、次のようになる。

$$\begin{aligned} \frac{a-b}{b-c} &= \frac{a}{c} \\ a(a-b) &= a(b-c) \\ a-b &= b-c \\ 2b &= a+c \\ b &= \frac{a+c}{2} \end{aligned}$$

よって、「算術的比例」が成立する3数についてはその中項 b は相加平均になる。これは「算術平均」とも言われる。

また、②式についても同様に b について解けば、次のようになる。

$$\begin{aligned} \frac{a-b}{b-c} &= \frac{b}{c} \\ c(a-b) &= b(b-c) \\ ca - bc &= b^2 - bc \\ b^2 &= ca \\ b &= \sqrt{ca} \end{aligned}$$

これにより、「幾何学的比例」が成立する3数についてはその中項 b は相乗平均になることが分かる。これは「幾何平均」とも言われる。②'についても同様の結果となる。

また、「調和的比例」にかんする中項 b は次のようにすれば、もとめることができる。

$$\begin{aligned} \frac{a-b}{b-c} &= \frac{a}{c} \\ c(a-b) &= a(b-c) \\ ca - cb &= ab - ca \\ b(c+a) &= 2ca \\ b &= \frac{2ca}{c+a} \\ &= \frac{2ca \cdot \frac{1}{ca}}{(c+a) \cdot \frac{1}{ca}} \\ &= \frac{2}{\frac{c+a}{ca}} \\ &= \frac{2}{\frac{1}{a} + \frac{1}{c}} \\ &= \frac{1}{\frac{1}{2} \left(\frac{1}{a} + \frac{1}{c} \right)} \end{aligned}$$

これはもとの系列を構成する a と c の逆数の相加平均であり、したがって、「調和的比例」が成立するとき、その関係を満たす3数の中間にある b は「調和平均」になる。

最後に、系列を構成する3数が、④式で表現されるような「反調和的比例」の関係にあるとき、その中項 b は「反調和平均」 M となる。 M は④式から次のように誘導されるからである (なお、④'からも同様の結果を得る)。

$$\begin{aligned} \frac{b-c}{a-b} &= \frac{a}{c} \\ c(b-c) &= a(a-b) \\ bc - c^2 &= a^2 - ab \\ b(a+c) &= a^2 + c^2 \\ b &= \frac{a^2 + c^2}{a+c} \end{aligned}$$

で定義される。 $a_1=+4$, $a_2=-3$ のとき、これらの値を(4)式に代入すれば、 $M=25$ となる。この $M=25$ は、大きい項の値(+4)よりも大きい。しかし、25という値はキズィーニの定義式である(3)式を2数に限定したときの一般式

$$f(M, M) = f(a_1, a_2) \quad (3)'$$

の系として誘導される反調和平均の計算式((4)式)を満たしている。

これにたいして、ジニはコーシーの定義を基本に据えて、「平均」とは、大きい値の項よりも小さくなく、かつ、小さい値の項よりも小さくない項の値であると規定している。このような「平均」は小さい値の項と大きい値の項の間にあることから、ジニは「平均」が「内部性の要請 (il requisito della internalità)」もしくは「内部性の条件 (la condizione della internalità)」を満たすと言っている (57)。

$(a_1, a_2) = (+4, -3)$ という組にキズィーニの定義式を当てはめてもとめた $M=25$ は、上に述べた「内部性の要請」を満たさない。このために、たとえ「解析的平均」と見なしうる数値であろうとも、「内部性の要請」を満たさない数値は、これを平均とは見なしがたいとジニは考えた。これがキズィーニをジニが批判したときの第2の理由である。

このように「内部性の要請」は平均概念を考察するときに重要な役割を果たしている。この「要請」は、コーシーの定義とその趣旨において同様である。しかし、原系列が実数で構成され、たとえ

$$f(M, M, \dots, M) = f(x_1, x_2, \dots, x_n) \quad (3)[再掲]$$

を満たしていようとも、 M が虚数となる場合には、平均値たり得ないことなどを条件に

していることに注目すると、ジニは、コーシーの定義をさらに厳密に規定しようと企図したとすることができる。

ジニは(3)式を満たす M であっても、「解析的平均」たり得ないことがあることを重く見て、この M を一般に「平衡数 (adeguati numerici: 直訳すれば「数的平衡」)」と名づけ、そのなかでとくに「内部性の要請」を満たす実数を「解析的平均」と規定している (134 f.)。「平衡数」について述べた箇所でジニは相加平均、相乗平均、累乗平均(平方平均はもっとも単純で特殊な累乗平均と規定される)などの場合には、「平衡数」が平均に一致することを指摘している。この限りでは、キズィーニの定義にもとづいて算出される数量は、平均とすることができる。しかし、「平衡数」がつねに「解析的平均」であるとは言いがたい場合があることを挙げて、キズィーニを批判した。

このようなジニの平均概念は、ピュタゴラス派の数学理論の研究に淵源すると考えられる。ピュタゴラス派の検討については今後の課題とすることにして、さしあたり、ここでは、「平均」と訳される単語を手がかりにして、キズィーニとジニの対立を考えてみる。

平均を表す英語には、averageとmeanがある。averageはアラビア語のawāriyaを語源とする。このアラビア語には「傷物」という意味がある。損害保険の分野では、これが転じて航海で発生した損害(「海損」)を意味するようになり、また「損害額を割り当てる」という動詞としても使用されるようになった。そこからさらに、averageは、さまざまな値を一様に「ならす(平す、均す)」という意味での「平均」として使用されるようになったと考えられている。この点でaverageという言葉は、さまざまな値の項からなる系列 x_1, x_2, \dots, x_n を「平し(均し)」て、単一の x によって諸項の値の代替とみなし、それによって系列を代表させるという、(ジ

ニの意味ではない) いわゆる「計算的平均」の代替機能 (あるいは代表機能) を表現している。キズィーニによる平均の定義式はこの意味で average に対応している。

これにたいして、コーシーやジニの平均には mean が該当する。mean の語源はラテン語の medianus (中間にあるもの) である (イタリア語の media[sing.], medie[pl.] も同様であり、綴りはこちらの方が語源に近い)。このために、英語の mean には、「両極端の中央に位置するもの」という意味もあり、さらには意図 (目的, 目標) と結果 (到達点) との間であってそれらを結びつける「手段」「方法」という意味や生きるための「手だて」(生活物資) を購入するための「収入」や「財産」という意味がある。ジニの「内部性の要請」は mean の語源としての medianus に対応している。この点で、ジニとキズィーニの見解の相違は medianus 派と awāriya 派との対立と見ることができる。

む す び

本稿では、ジニが『平均論』で展開した考察の根幹をなす平均の定義を取り上げた。その検討結果は以下のように要約される。

- (1) 著書『平均論』はジニの単著と言うよりは、第 2 次世界大戦以前からジニが暖めてきた構想を実現すべくローマ大学で組織された研究会の成果であり、ジニ学派において共通理解に至った見解をとりまとめた著作と見なすことができる。
- (2) ジニは平均を「解析的平均」と「非解析的平均」に 2 分した。このうち、後者の「非解析的平均」は「位置上の平均」とも言われ、今日、われわれが「計算的平均」と対照させる「位置上の平均」とはその内容に異なるところがない。
- (3) ジニの「解析的平均」は平均をもとめ

るべき実数の系列にかんする関数関係において把握される。この点が、関数関係によっては表現できない「非解析的平均」とは異なっている。

- (4) ジニの「計算的平均」とは、平均をもとめるべき系列のなかには現実に対応する項を見出し得ない数値のことである。このため、われわれが通常、用いる「計算的平均」とは異なった意味をもっている。
- (5) 概して、ジニの「解析的平均」には「固定的」、非解析的平均には「弛緩的」という特徴がある。ここに「固定的」とは系列を構成する項の数量的規定性にかんする変化に平均の値が鋭敏に反応することを言う。
- (6) キズィーニは、平均 M が関数関係

$$f(M, M, \dots, M) = f(x_1, x_2, \dots, x_n) \quad (3)[再掲]$$

を満足する値であると定義した。

- (7) (3)式を満たす M をジニは「平衡数」と名づけた。相加平均や相乗平均などは「平衡数」と一致するが、すべての「平衡数」がジニの「平均」と一致するとは限らない。このために、(3)式は「解析的平均」の一般式として不適切であるとジニは述べた。
- (8) (3)式が「非解析的平均」はもとより、「解析的平均」の定義式としても有効でないとジニが主張した根拠は「内部性の要請」である。この「要請」は、平均が、実数の系列において、小さい値の項よりも小さくなく、大きい値の項よりも大きくないという、あらゆる平均が満たすべき「条件」となっている。
- (9) 本稿では、(3)式をもって平均の定義式とするキズィーニを awāriya 派に属す論者に分類した。これにたいして、「内

部性の要請」を満たす実数をもって平均と見なすジニ(学派)を medianus 派と名づけて、両者を対照させた。

「解析的平均」と「非解析的平均」とを問わず、いずれの平均であろうとも遵守すべきとされる「内部性の要請」は、ピュタゴラス

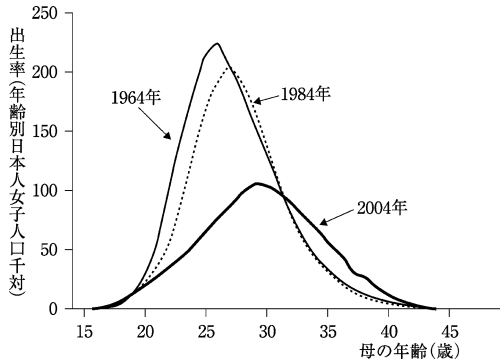
学派による比例関係の研究を踏まえた1つの結論である。したがって、ピュタゴラス学派の見解をジニはどのように見ていたのかを理解することによって、ジニ『平均論』の根底を流れる思考様式を検討することができる。この点の考察は今後の課題である。

(数学注) メディアン(中央値) Me と
モード(最頻値) Mo

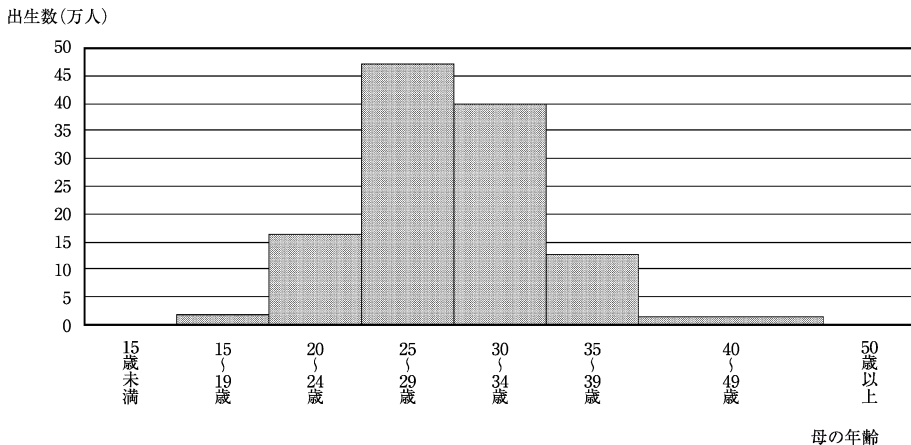
付図1の縦軸は出生率(年齢別日本人女子人口千対)を示している。この図から母の出産年齢は上昇傾向にあることが分かる。付図1では、縦軸に相対数がとられているが、出生数(絶対数)をとることも可能である(付図2参照)。そのような度数分布図において、出産した母の総人数を半分に分かつ年齢をメディアン(中央値) Me という。また、もっとも出産数が多い母の年齢をモード(最頻値) Mo という。

この国では相加平均や相乗平均が「計算的平均」と言われるのにたいして、 Me や Mo は「位置上の平均」と言われている。記述統計学が比較的重視されていた戦前・戦中期には、統計学の教科書では「計算的」と「位置上」という2種類の平均が取り上げられていたが、戦後に

なってきたら確率論基調の統計理論(推測統計学)が主流になり、いつの間にか「位置上の平均」は忘れ去られたかの感を呈していた。



付図1 母の年齢別にみた出生率の年次比較
(出所)『国民衛生の動向・厚生指標』(臨時増刊) 第53巻第9号 2006年, p.41。



付図2 母の年齢階級別出生数の度数分布図

付表1 母の年齢階級別出生数* (2000年) (人)

15歳未満	15～19	20～24	25～29	30～34	35～39	40～49	50歳以上
43	19,729	161,361	470,833	396,901	126,409	14,848	6

*総数 (年齢不詳を含む) 1,190,547人

*総数 (年齢不詳を除く) 1,190,130人

(出所) 総務省統計局『日本の統計 2006年版』日本統計協会, 2007年, p.26より抜粋。

ところが、近年、OECDがメディアンに注目するようになった。OECDの調査研究報告書では、所得分布のメディアンに該当する所得の2分の1以下の所得層を「貧困層」と規定し、全体に占めるその割合(「相対的貧困率」)をもとめ、それによる国際比較が試みられている。経済格差が社会問題になるにつれて「位置上の平均」は、その分析手法として復位したかに見える。メディアンとモードにかんするこの「数学注」は、内容的には旧聞に属し、しかも初等数学的な叙述に終始するので、屋上屋を架するとの「そしり」を受けるかもしれない。しかし、このような現状にあっては、幾ばくかの有用性があるのではないかと考えて、先学の叙述を参考にして、数値例を付けた数学注をおくことにした。

1. メディアン (中央値)

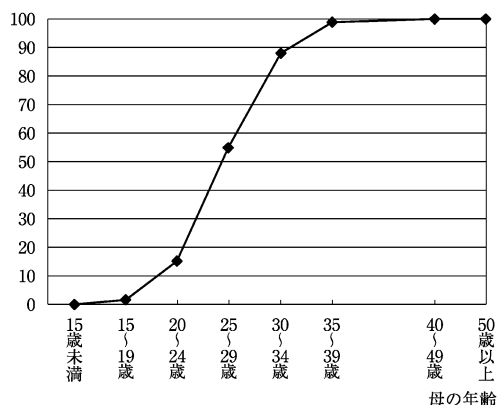
付表1は2000年における母親の年齢階級別出生数(日本)を表章している。

付表1から度数分布図を描くと付図2のようになる。

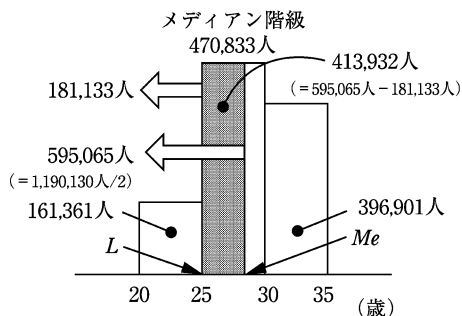
メディアン Me とは、度数分布の総度数を半分にする横軸の座標であたえられる。 Me がどのあたりにあるかは、累積相対度数を折れ線グラフに書いて、縦軸の50%に該当する階級の年齢を読み取れば、おおよその見当がつく。そのために付表1から累積相対度数のグラフを描くことにする(付図3)。

付図3から、出生総数の1/2(累積相対度数が50%)に当たる母の年齢(Me)は、「25～29歳」階級に落ちていることが分かる。この階級を「メディアン階級」と言う。ところが、「メディアン階級」が分かっても、何歳の母を境にして、出生数が出生総数の半分になっているかは分からない。

累積出生相対度数(%)



付図3 母の年齢階級別累積出生相対度数 (2000年)



付図4 メディアン階級 (25～29歳) を中心とする度数分布図 (部分)

そこで、 Me の値を特定するために、付図2から、メディアン階級を真ん中にして、前後の階級(全部で3つの階級)を抜き出す(付図4)。この付図4には、メディアン階級(階級間隔 [= 階級幅] は5歳、これを c とおく)の人数(度数) f が470,833人であると記載されている。

ここで、メディアンの定義を想起する。それによれば、年齢が Me までの母の人数は、す

すべての母の人数(総度数, ただし年齢不詳を除く) 1,190,130人 (N) の半分, すなわち $N/2$ 人 (=595,065人) である。この $N/2$ は, ①一番若い母の年齢からメディアン階級の下限 L までの母の人数 F と②メディアン階級の下限 L からメディアン Me までの人数に分解される(付図4参照)。

このことをメディアン階級に着目して考えてみると次のようになる。すなわち, ①総度数の半分 $N/2$ 人 (=595,065人) から②25歳未満までの階級(メディアン階級の左側にあるすべての階級)に属す母の度数 F (181,133人) を引けば ($N/2 - F$), メディアン階級に属す母親のうち, 付図4の L から Me までの間にいる人数が得られる。その人数は 413,932人 (=595,065人 - 181,133人) である。メディアン階級の下限(25歳)を L で表すとき(付図4), ($Me - L$)で示される範囲には ($N/2 - F$)人の母親がいることになる。このことから次のことが分かる。

- ① メディアン階級の階級間隔 c には, メディアン階級の度数 f が対応していること。
- ② メディアン階級の下限 L からメディアン Me までの間隔($Me - L$)には, 総度数の半分 ($N/2$)に足りない度数をメディアン階級から補う度数($N/2 - F$)が対応していること。ここに, F はメディアン階級よりも下位の全階級の度数である

①と②で述べた関係を数式で表現すれば, 次のようになる。

$$Me = L + \frac{\frac{N}{2} - F}{f} \times c \tag{ii}$$

これに関連数値を代入すれば,

$$Me = 25 + \frac{1,190,130}{2} - 181,133}{470,833} \times 5 = 29.40(\text{歳})$$

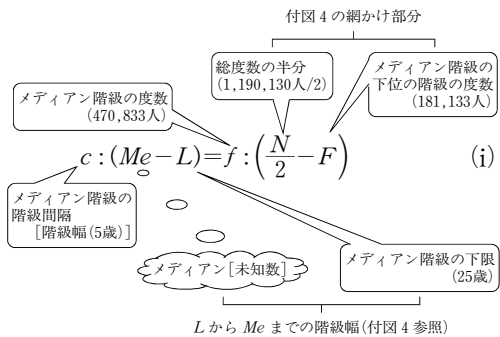
となり, メディアンは 29.4歳である。

以上から, メディアンは, それが存在する階級の幅(メディアン階級の階級間隔: c)をその階級の下限 L からメディアン Me までの度数で按分していることが分かる。

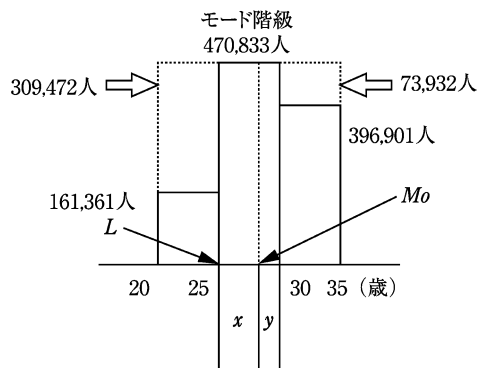
2. モード(最頻値)

モード Mo は, 度数分布においてもっとも度数が大きい値である。付表1(付図2)から明らかのように, 最大度数の階級(モード階級)は「25~29歳」階級である。モード階級の下限(25歳)とその右隣の階級の下限(30歳)の相加平均(27.5歳)をモードの値 Mo と見なすことがある。近似的にはこれでもよいが, これは幾分, 厳密さに欠ける。

モードの数値を特定する目的で, 付図2からモード階級とその前後の階級にかんする度数分布図を抜き出すことにする(付図5)。モード階級を中心とする3つの階級分布が付図5のよう



(i) 式を整理すると, 次式を得る。



付図5 モードは大きい度数の階級(30~34歳階級)に引き寄せられる

になるとき、モード階級の右隣に位置する階級（「30～34 歳」階級）の度数（396,901 人）は、「20～24 歳」階級（モード階級の左隣の階級）の度数（161,361 人）よりも多い。この場合、モード Mo は度数が大きい右隣の階級に引き寄せられていると考えるのが自然である。すなわち、付図 5 に示すように $x > y$ である（これとは逆に、モード階級の左隣にある階級の度数が右隣の階級の度数よりも大きい場合には、モードは左隣の階級に引き寄せられる）。

Mo の数値を特定するためには、モード階級の下限 L （25 歳）から Mo がどれだけ乖離しているか、換言すれば Mo の位置を規定する x がどれだけの大きさであるかを特定しなければならない。 x の値がもとめられれば、モード階級の下限 L にこの x を足すことによって、 Mo が分かる ($Mo = L + x$)。

さて、 x によって Mo の値を特定するには、付図 5 において $x > y$ という大小関係にある x (と y) が、モード階級とその両隣の階級との度数差によって規定されると考えればよい（米澤治文・一条勝夫『講要 統計学』日本評論社、1958 年、p.50、および足利末男『社会統計学の基礎』晃洋書房 1982 年、p.142 f. 参照）。

付図 5 にもとづいてモード階級の両隣の階級について関連数値を付表 2 に表章する。

この度数差の割合の違いが x の値を規定し、その x によってモードの位置が定まると考えることが、モードの数値的特定におけるポイントである。度数差の割合が小さい階級（実際の度数がモード階級の度数により近い階級）ほど、モードをその階級に引き寄せる。そして、逆に、度数差の割合が大きい階級（実際の度数がモード階級の度数からより大きく乖離している階級）ほど、

モードをその階級から遠ざけている。このことは付図 5 から直観的に理解できる。

設例では、度数差の割合がより大きい階級は、モード階級の左隣に位置する階級（「20～24 歳」階級）である。この階級に着目すれば、モード Mo の値は、この階級に x だけ近づいていると言うよりは、度数差の割合がより小さい階級（「30～34 歳」階級）の方へと（モード階級の下限 L （25 歳）から距離にして x だけ右方に）「押しやる」と見るべきであろう。

以上の考察によって、モード階級の階級間隔（5 歳）($=x+y$)——これを c とおく——を度数差の割合で按分すれば、モード Mo が定まることになる（付図 5、付表 2 参照）。すなわち

$$\begin{aligned} x &: (x+y) \\ &= \frac{309,472}{309,472 + 73,932} : \left(\frac{309,472}{309,472 + 73,932} \right. \\ &\quad \left. + \frac{73,932}{309,472 + 73,932} \right) \quad (\text{iii}) \end{aligned}$$

を満たす x をもとめれば、 Mo の位置を定めることができる。上で述べたように $x+y$ はモード階級の階級間隔 c であるから、 $x+y=5$ ($=c$) になる。これを (iii) 式に代入して整理すれば、

$$\begin{aligned} x &: 5 = \frac{309,472}{383,404} : 1 \\ \therefore x &= 5 \times \frac{309,472}{383,404} \\ &= 4.04 \quad (\text{iv}) \end{aligned}$$

を得る。これにより付図 5 の x は $x=4.0$ になる。モード階級（「25～29 歳」階級）の下限 (L) は 25 であるから、

付表 2 モード階級の左右の階級別度数差とその割合

階 級	左側の階級 (20～24 歳)	右側の階級 (30～34 歳)
モード階級との度数差	309,472 人 (=470,833-161,361)	73,932 人 (=470,833-396,901)
度数差の割合	$\frac{309,472}{383,404} \left(= \frac{309,472}{309,472 + 73,932} \right)$	$\frac{73,932}{383,404} \left(= \frac{73,932}{309,472 + 73,932} \right)$

$$\begin{aligned}
 Mo &= L + x \\
 &= 25 + 4.0 \\
 &= 29.0
 \end{aligned}
 \tag{v}$$

となり、モードは29.0歳である。

(iv)式と(v)式を参考にすれば、モード Mo をもとめる一般式は次のようになる。

$$Mo = L + c \times \frac{D_l}{D_l + D_r}
 \tag{vi}$$

ただし、 L はモード階級の下限。

c はモード階級の階級間隔。

D_l はモード階級の左側の階級にかんする度数差。

D_r はモード階級の右側の階級にかんする度数差。

ここで、モード階級の度数を f 、その左右に隣接する階級の度数をそれぞれ f_l および f_r とすれば、それぞれの階級の度数差は

$$\begin{aligned}
 D_l &= f - f_l \\
 D_r &= f - f_r
 \end{aligned}$$

である。この関係を(vi)式に代入すれば、モード Mo をもとめる一般式は次のようにも表現することができる。

$$Mo = L + c \times \frac{f - f_l}{(f - f_l) + (f - f_r)}
 \tag{vii}$$

$$= L + c \times \frac{f - f_l}{2f - (f_l + f_r)}
 \tag{vii}'$$

この(vii)式(または(vii)')式を用いれば、モード階級およびそれを挟む階級の度数から Mo をもとめることができる。モード階級の左隣りと右隣りのいずれかに階級が存在しないとき、あるいは、左端または右端の階級がモード階級であるときには、適宜、 $f_l = 0$ または $f_r = 0$ を(vii)式に代入すれば、 Mo が計算される。

3. メディアン、モード、相加平均の間の数学的関係 — ドゥードソンの近似式 —

中程度 (a moderate degree) の非対称分布にあってはメディアン Me 、モード Mo 、相加平均 \bar{x} の間には

$$|Mo - Me| = \frac{2}{3} \times |Mo - \bar{x}|
 \tag{viii}$$

という数学的関係があると考えられていた時期があった。カール・ピアソンはこの関係を特殊な分布について証明した (Pearson, Karl, "Contribution to the Mathematical Theory of Evolution, II. Skew Variation in Homogeneous Material," *Philosophical Transactions of the Royal Society of London*, Ser. A., Vol. 187, 1895; also in *Karl Pearson's Early Statistical Papers*, Cambridge 1948.)。

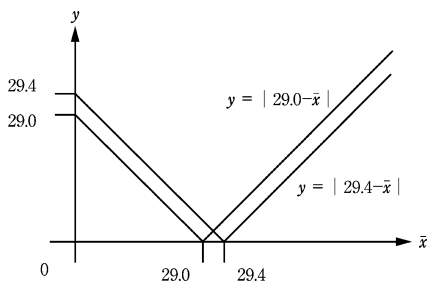
これにたいして、ドゥードソン (Arthur T. Doodson) は、分布型や分布関数のパラメータの値によって近似度が異なるが、非対称分布の Me 、 Mo 、 \bar{x} については近似的に

$$|Me - \bar{x}| = \frac{2}{3} \times |Mo - \bar{x}|
 \tag{ix}$$

が成り立つことを証明した (Doodson, Arthur T., "Relation of the Mode, Median and Mean in Frequency Curves," *Biometrika*, Vol. XI, 1915-17, pp. 425ff.). この数学注の1.と2.で取り上げた母の年齢分布にかんする相加平均 \bar{x} の近似値は、この(ix)式によって、もとめることができる。1.から $Me = 29.4$ 歳、また2.から $Mo = 29.0$ 歳となった。この数値を(ix)式に代入すれば、

$$|29.4 - \bar{x}| = \frac{2}{3} \times |29.0 - \bar{x}|$$

となる。したがって、もとめる \bar{x} の値は



付図6 Me, Mo, \bar{x} の近似的関係

$$\begin{cases} y = |29.4 - \bar{x}| \\ y = |29.0 - \bar{x}| \end{cases}$$

を満たす。この方程式をグラフで表示すれば、付図6のようになる。

付図6から、もとめる解は次の連立方程式の解としてあたえられることが分かる。すなわち、

$$\begin{cases} y = 29.4 - \bar{x} \\ y = \bar{x} - 29.0 \end{cases}$$

これを解けば

$$29.4 - \bar{x} = \bar{x} - 29.0$$

$$2\bar{x} = 58.4$$

$$\therefore \bar{x} = 29.2(\text{歳})$$

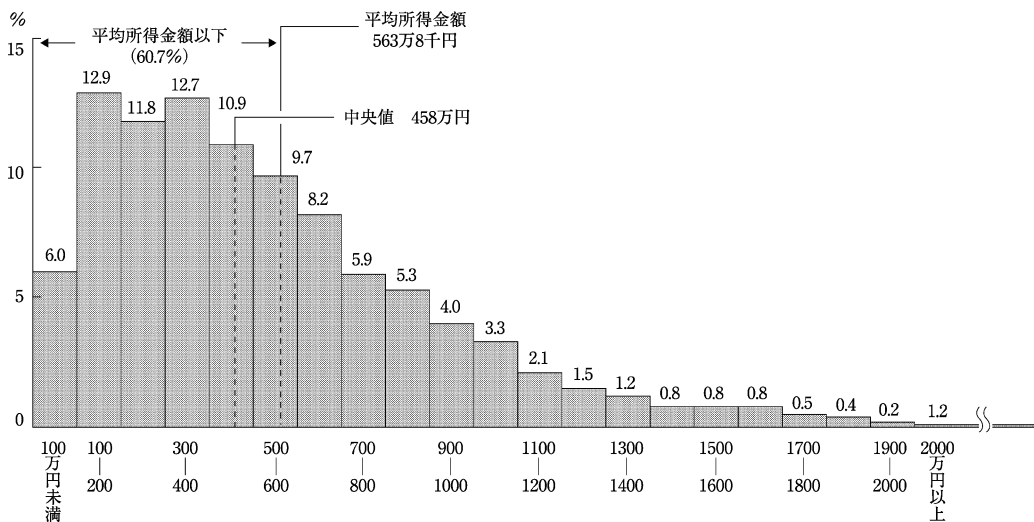
となつて、母の平均年齢(相加平均)の近似値は29.2歳となる。

それでは付図7のように、所得の総平均(相加平均、厳密には加重相加平均) \bar{x} とメディアン Me があたえられているときは、モード Mo の近似値はどうなるであろうか。

$Me = 458$ 万円, $\bar{x} = 563.8$ 万円であるから, $Me < \bar{x}$ である。また、付図7から明らかのようにモード階級は総平均が存在する階級よりも小さい階級にあるので、 $Mo < \bar{x}$ である。したがって、このような場合には、ドゥードソンの近似式((ix)式)は、

$$\bar{x} - Me = \frac{2}{3} \times (\bar{x} - Mo)$$

である。この式に関連数値を代入すれば、



付図7 所得分布 (2006年調査)

(出所) 厚生労働省『国民生活基礎調査』(2006年)

<http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa06/2-1.html>, accessed on March 3, 2007.

$$563.8 - 458 = \frac{2}{3} \times (458 - M_o)$$

$$\frac{2}{3} M_o = -(563.8 - 458) + \frac{2}{3} \times 458$$

$$\therefore M_o = 299.3 \text{ (万円)}$$

となり、モード M_o の近似値は 299 万 3000 円となる。

ここで、ドゥードソンの近似式の適用結果と比較する目的から、ピアソンの近似式 ((viii) 式) に関連数値 ($M_e = 458$ 万円, $\bar{x} = 563.8$ 万円) を代入すると、

$$|M_o - 458| = \frac{2}{3} \times |M_o - 563.8|$$

を得る。上式の辺々を二乗すると、

$$(M_o - 458)^2 = \left(\frac{2}{3}\right)^2 \times (M_o - 563.8)^2$$

になる。これを整理すると

$$\frac{4}{9} M_o^2 - 414.78 M_o - 527634.44 = 0$$

であり、その解は $M_o = 1,651.9$, -718.7 である。題意より、もとめるモードは 1651 万 9000 円である。この値とドゥードソンの近似式があたえる値 (299.3 万円) を比較すると、後者の方が良好な近似値をあたえていることが分かる。