《論説》

平均差とジーニ係数

木　村　和　範

もくじ

はじめに

ジーニ係数の数理的意味を説明するために，ローレンツ曲線が用いられることがある。この係数を G とおけば，G は，ローレンツ曲線と均等分布直線に囲まれた三日月形の図形の面積（ジーニのいわゆる「集中面積」λ）と 1/2（直角二等辺三角形の面積）との比率として定義され，

$$G=\frac{\lambda}{\frac{1}{2}} \tag{1}$$

という数学的関係にあるという説明がそれである。この説明は，図形という視覚的手段を用いているために，ジーニ係数の意味の具体的な理解を可能にする。

また，(1)式を

$$G=2\lambda \tag{1'}$$

と変形し，これを用いて，ジーニ係数は集中面積の 2 倍をあたえるとする定義もある。ジーニ係数が「集中比（rapporto di concentrazione）」R と言われていたころ，すでに(1')式によって R を定義する試みがあった[1)]。しかし，このような説明では，$\lambda=0$ という完全平等分布（図形では直角二等辺三角形の面積［1/2］としてあたえられる）を基準にして，ジーニ係数 G が（不平等の程度を示す）集中面積λを計測しているという，G に内在する特質ばかりか，G が集中比と

1）Pietra, Gaetano, “Delle relazioni tra gli indici di variabilità,” *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti*, Anno accademico 1914–15, Tomo LXXIV, Parte seconda, p.780.［Pietra (1915)］

いう名の比率であるという性質までもが見失われてしまう。その意味では集中比の考案者であるジーニに従って，「ジニ係数はローレンツ曲線と対角線 OB に囲まれた部分の面積……を三角形 OAB の面積で除したものに等しい」[2] と見るのが至当である。

ジーニ係数の数理的意味を説明するとき，平均差 $\varDelta$ が用いられることもある。ここに言う平均差とは，系列を構成する個体の数量的特性（ジーニのいわゆる強度［intensità］）を 2 つずつ組み合わせて，それについて可能な差をすべて計算し，その差の総和 T を差の個数 $N[=n(n-1)]$ で割ったときに得られる商のことである。この平均差 $\varDelta$ を用いると，ジーニ係数 G は，$\varDelta$ を被除数，強度の相加平均 M_n の 2 倍を除数とする統計量

$$G=\frac{\varDelta}{2M_n} \tag{2}$$

と定義される。⑵式を用いたジーニみずからによる説明は，コールズ委員会（アメリカ）の研究集録（1936 年）[3] でも見られる。しかし，そのような説明は，ジーニ係数が集中比 R として初めて定式化された 1914 年論文においてすでに試みられている[4]。

これまでは，パレートの所得分布研究に始まりローレンツを経てジーニにいたる，その理論的展開過程を明らかにする一環として，集中比の定義式からその計算式がどのように誘導されるかを検討するとともに，その集中比 R とローレンツ曲線との関係を調べてきた。しかし，それらの検討では，ジーニの 1914 年論文における主要論点のうちの少なくともひとつは取り上げられていなかった。それは，集中比と平均差の数学的関係である。本稿では，ジーニが変動性指数として位置づけた平均差を集中比との関係に限定して取り上げ，1914 年論文にたいするこれまでの検討を補完したい。

1．変動性指数と強度差

⑴　変動性指数

ジーニは，「変動性指数（indici di variabilità）」にかんする論文「変動性と変化性 ── 分布と統計的関係の研究のために ──」（1912 年）を執筆し，平均差が，統計系列の構成要素の変動性（variabilità）を計測するための測度（変動性指数）として有効であると主張した[5]。

そこで，上記 1912 年論文でジーニが定式化した平均差 $\varDelta$ を取り上げるに先だって，変動性指数の意義にかんする彼の見解を見ておく。

ジーニは系列を構成する要素（「特性［carrateri］」）の「様相（modalità）」を質的と量的に分類した（p.3[6]）。「様相」の違いが量的に示される場合には，その「特性」が「変動する（variare）」と言う。質的な違いが問題となる場合には，その「特性」が「変

2）高山憲之「富と所得の分布」『経済学大辞典（第 2 版）』Ⅰ 東洋経済新報社　1980 年，p.473。石田望「所得配分不均等度の測定 ── ジニ係数適用拡大の試み ──」『東京経大学会誌』第 146 号 1986 年，p.116 も参照。

3）Gini, C., “On the Measurement of Concentration with Special Reference to Income and Wealth,” *Abstracts of Papers presented at the Cowles Commission Research Conference on Economics and Statistics,* Colorado 1936, p.77.

4）Gini, C., “Sulla misura della concentrazione e delle variabilità dei carratteri,” *Atti del Reale Istituto di Scienze, Lettere ed Arti,* Tomo LXXIII, Parte seconda, 1913-14, pp.1236ff.［Gini (1914)］

5）Gini, C., “Variabilità e Mutabilità, contributo allo studio delle distribuzioni e delle relazione statistiche,” *Studi Economico-Giuridici delle R. Università di Cagliari,* Vol.3, Parte seconda, 1912.［Gini (1912)］

6）以下，とくに断らない限り，本文の（　）内に示した数字は，Gini（1912）のページを示す。

化する（mutare）」と言う（p.15）。そして，「変動」を計測するための指標を「変動性指数（indici di variabilità）」と名づけた。他方で「変化」を計測するための指標を「変化性指数（indici di mutabilità）」と名づけた。

ジーニによれば，変動性指数は2つのカテゴリーに分かれる。第1カテゴリーに属す変動性指数は「特性にかんする測定値が実際の大きさ（grandezza effettiva）［真値］からどの程度乖離しているか」という問いにたいする数値的な解答をあたえる。n個の観測値と真値との乖離をe_iとし，mを1以上の整数とする。このとき，

$$
{}^{m}S_A=\sqrt[m]{\frac{\sum|e_i|^m}{n}} \tag{3}
$$

は第1カテゴリーの変動性指数にかんする一般式である（p.18f.）。天文学や測地学（測量学）では，取り扱う観測誤差が確率的に正負の値をとる偶発誤差と考えられている。誤差論（観測値結合論）はこのような前提をおいて，観測値の相加平均を真値の「本当らしい値（il valore probabile）」と見なした。そして，個別観測値とその相加平均との乖離（平均偏差）や観測値の散布を示すために(3)式を基本形とする多数の尺度が考案された（p. 18）。

これにたいして，第2カテゴリーに属す変動性指数は，「多様な実際の大きさ（le diverse grandezza effettive）がそれらどうしで，どれだけ異なっているか」を取り扱う。確かに，第1カテゴリーに属す変動性指数は，それとしてその有効性を否定できないが，「目的上の相違」は「方法の多様性」を要請するのであって，人口論，人類学，生物学，経済学では，第1カテゴリーとは異なった変動性指数が必要とされる（p.20）。人口論などの分野で，「それらの特性にかんする変動性の測定という目的に適合しているのは，断じて算術平均［相加平均］からの測定量の乖離の強度を測定する指標ではなくて，実測値間の差異の強度を計測する指標なのである」（p.20）。ジーニはこのための「指標」を第2カテゴリーに属す変動性指数と位置づけて，1912年論文でその解明を試みた。

集中比R（ジーニ係数）を初めて定式化した論文（1914年）[7]を見れば，ジーニは，第1カテゴリーに属す変動性指数（相加平均，分散，平均偏差など）が，富，資産，遺産，（住居の）賃貸料，出生数，婚姻数，死亡数など，特性間の差を問題とする「経済学や人口論にかんする統計研究」で十分にその機能を一般的に果たすかどうかにたいして懐疑的であったことが分かる。そして，ジーニは，1914年論文では，1912年論文で定式化した第2カテゴリーに属す変動性指数としての「平均差」を取り上げ，これと集中比との数学的関連を明らかにした。ただし，1914年論文では，平均差が誘導済みの統計量であることから，その計算式が前提となっている。そこで，1914年論文において集中比を定義するときに用いられた平均差$\varDelta$が1912年論文ではどのように誘導されているかを次項以降で追跡する。

(2) 強度差の総和

1912年論文（p.22）で誘導した平均差$\varDelta$

$$
\varDelta=\frac{2}{n(n-1)}\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i) \tag{4}
$$

ただし，a_i：系列を構成する各項にかんする強度（intensità）（以下，「強度」は各項の数量的規定性を意味する）。

i：強度を昇順に並べたときの順位。

n：系列を構成する項数（強度の個数）。

$\varDelta$：$n(n-1)$個ある$a_j-a_i \geqq 0\,(i<j)$の相加平均（平均差）。

7）Gini (1914), p.1237.

を用いてジーニは集中比 R を定義している。ジーニにならって(4)式を誘導する目的で，単一の統計系列を構成する項の強度を 2 つずつ取り出し，それを組合せることにする。そして，2 つの強度の差がいずれも非負であたえられるようにするために，小さくない強度を被減数，大きくない強度を減数としたときに得られる強度差を考える。ジーニは，後に明らかになる理由（本節(3)参照）から，同一の項の強度どうしについても組を作り，その差をもとめている。そして，彼は，これらの差を一般に「n 個の諸量の差（differenza tra n quantità）」（p.20）と言っている。事柄をより明確にするために，本稿ではこの差を「強度差（differenza d'intensità, difference of intensity）」と言うことにする。この「強度差」という言葉は，19 世紀後半における誤差論研究者 W. ヨルダンが，2 つの観測値どうしの差を「観測差（Beobachtungsdifferenz）」[8) と言っていることに着想を得た筆者の造語である。強度差の代わりに観測差と言うことも可能である。しかし，ここではジーニの「強度」という概念にこだわることにした。誤差論における観測値とジーニが集中比でその分布を計測しようとした強度とでは，その数学的性質が異なるので，両者の違いを明確にする必要があるからである。この違いは，基本的には強度と観測値というそれぞれの変量の分布が正規分布に従うかどうかということにある。

さて，ジーニによれば，(4)式であたえられる平均差 $\varDelta$ は，強度差の総和 T を強度差の個数 N で割ったときの商に等しい。したがって，$\varDelta$ をもとめるには（i）強度差の総和 T と（ii）その個数 N を知る必要がある。そこで，以下では 2 つの強度の差（強度差）にかんする表 1 ～表 6 を用いて，まず強度差の総和 T をもとめる。これらの表のいずれにおいても減数もしくは被減数のいずれか一方に必ず位置する強度がひとつだけある。たとえば，表 1 では a_1 がそれであり，表 2 では a_2 がそれである。表 3 以降についても同様である。このように，表ごとに強度差を見た場合，そこには共通して見られる強度 a_i がある。それは，ある表と他の表とを識別するマーカーとなっている。ジーニの用語ではないが，本稿では，そのような強度を「共通強度（intensità commune, common intensity）」と言うことにする。共通強度別に強度差の合計をもとめ，それをまとめて，表 1 ～表 6 に示した。

これらの表は，すべての可能な強度差を網羅しているので，表ごとに（共通強度別に）

表 1　強度差とその合計（その 1）

強度差（共通強度 a_1）	強度差の合計
$a_1 - a_1$ $a_2 - a_1$ $a_3 - a_1$ $\vdots$ $a_{n-2} - a_1$ $a_{n-1} - a_1$ $a_n - a_1$	$(a_1+a_2+a_3+\cdots+a_{n-2}+a_{n-1}+a_n)-na_1$ $=(a_1+a_2+a_3+\cdots+a_{n-2}+a_{n-1}+a_n-na_1)+(a_1-a_1)$
強度差の個数	n
同一の項どうしの強度差を除いたときの強度差の個数	$n-1$ （同一の項どうしの強度差は a_1-a_1）

（注）いずれの強度差も非負とする。$j>i$ のとき $a_j>a_i$。以下同様。

8) Jordan, W., "Ueber die Bestimmung der Genauigkeit mehrfach wiederholter Beobachtungen *einer* Unbekannten," *Astronomische Nachrichten,* Vol.74, 1869, p.217.［Jordan (1869)］

表２　強度差とその合計（その２）

強度差（共通強度 a_2）	強度差の合計
$a_2 - a_1$ $a_2 - a_2$ $a_3 - a_2$ $\vdots$ $a_{n-2} - a_2$ $a_{n-1} - a_2$ $a_n - a_2$	$\{a_2+a_3+\cdots+a_{n-2}+a_{n-1}+a_n-(n-1)a_2\}+\quad a_2-a_1$ $=\{a_2+a_3+\cdots+a_{n-2}+a_{n-1}+a_n-(n-1)a_2\}+(2a_2-a_1-a_2)$
強度差の個数	n
同一の項どうしの強度差を除いたときの強度差の個数	$n-1$ （同一の項どうしの強度差は a_2-a_2）

（注）強度差が非負となるようにするために，最初の組の強度差だけは共通強度 a_2 が被減数となっている。以下の表でも，共通強度が被減数となっている強度差があるのは，この表と同様に非負の強度差を得るためである。なお，差の絶対値をとれば，上の表のように a_2-a_1（一般に $a_j-a_i\,[j \geqq i]$）とする必要はなくなるが，本稿はジーニの所説を敷衍することが目的であるから，彼にならって強度差を $|a_i-a_j|$ とは表記しない。

表３　強度差とその合計（その３）

強度差（共通強度 a_3）	強度差の合計
$a_3 - a_1$ $a_3 - a_2$ $a_3 - a_3$ $\vdots$ $a_{n-2} - a_3$ $a_{n-1} - a_3$ $a_n - a_3$	$\{a_3+\cdots+a_{n-2}+a_{n-1}+a_n-(n-2)a_3\}+\{(a_3-a_1)+(a_3-a_2)\}$ $=\{a_3+\cdots+a_{n-2}+a_{n-1}+a_n-(n-2)a_3\}+(2a_3-a_1-a_2)$ $=\{a_3+\cdots+a_{n-2}+a_{n-1}+a_n-(n-2)a_3\}+(3a_3-a_1-a_2-a_3)$
強度差の個数	n
同一の項どうしの強度差を除いたときの強度差の個数	$n-1$ （同一の項どうしの強度差は a_3-a_3）

（注）この後，基準強度 a_4 以降の強度差については，紙幅の関係でその表を割愛し，共通強度 a_{n-2} 以降について上と同様の表を掲げることとする。

表４　強度差とその合計（その４）

強度差（共通強度 a_{n-2}）	強度差の合計
$a_{n-2} - a_1$ $a_{n-2} - a_2$ $\vdots$ $a_{n-2} - a_{n-3}$ $a_{n-2} - a_{n-2}$ $a_{n-1} - a_{n-2}$ $a_n - a_{n-2}$	$(a_{n-2}+a_{n-1}+a_n-3a_{n-2})+\ (n-3)a_{n-2}-a_1-a_2-\cdots-a_{n-3}$ $=(a_{n-2}+a_{n-1}+a_n-3a_{n-2})+\{(n-2)a_{n-2}-a_1-a_2-\cdots-a_{n-3}-a_{n-2}\}$
強度差の個数	n
同一の項どうしの強度差を除いたときの強度差の個数	$n-1$ （同一の項どうしの強度差は $a_{n-2}-a_{n-2}$）

表 5　強度差とその合計（その 5）

強度差（共通強度 a_{n-1}）	強度差の合計
$a_{n-1}-a_1$ $a_{n-1}-a_2$ $\vdots$ $a_{n-1}-a_{n-2}$ $a_{n-1}-a_{n-1}$ $a_n\quad -a_{n-1}$	$(a_{n-1}+a_n-2a_{n-1})+\ (n-2)a_{n-1}-a_1-a_2-\cdots-a_{n-2}$ $=(a_{n-1}+a_n-2a_{n-1})+\{(n-1)a_{n-1}-a_1-a_2-\cdots-a_{n-2}-a_{n-1}\}$
強度差の個数	n
同一の項どうしの強度差を 除いたときの強度差の個数	$n-1$ （同一の項どうしの強度差は $a_{n-1}-a_{n-1}$）

表 6　強度差とその合計（その 6）

強度差（共通強度 a_n）	強度差の合計
a_n-a_1 a_n-a_2 $\vdots$ a_n-a_{n-2} a_n-a_{n-1} a_n-a_n	$(a_n-a_n)+(n-1)a_n-a_1-a_2-\cdots\quad a_{n-2}-a_{n-1}$ $=(a_n-a_n)+\qquad (na_n-a_1-a_2-\cdots-a_{n-2}-a_{n-1}-a_n)$
強度差の個数	n
同一の項どうしの強度差を 除いたときの強度差の個数	$n-1$ （同一の項どうしの強度差は a_n-a_n）

計算された強度差の合計をすべて合算すれば，その総和 T をもとめることができる。そのために，共通強度別の強度差の合計を表の順に再掲する（表 7）。

表 7 に記載した共通強度別強度差合計の総和 T は次のようにすればもとめることができる（表 7 最下欄参照）。

$$
\begin{aligned}
T=&\{a_1+2a_2+3a_3+\cdots\\
&+(n-2)a_{n-2}+(n-1)a_{n-1}+na_n\}\\
&-\{na_1+(n-1)a_2+(n-2)a_3+\cdots\\
&+3a_{n-2}+2a_{n-1}+a_n\}\\
&+\{a_1+2a_2+3a_3+\cdots\\
&+(n-2)a_{n-2}+(n-1)a_{n-1}+na_n\}\\
&-\{na_1+(n-1)a_2+(n-2)a_3+\cdots\\
&+3a_{n-2}+2a_{n-1}+a_n\}\\
=&2\{a_1+2a_2+3a_3+\cdots\\
&+(n-2)a_{n-2}+(n-1)a_{n-1}+na_n\}\\
&-2\{na_1+(n-1)a_2+(n-2)a_3+\cdots\\
&+3a_{n-2}+2a_{n-1}+a_n\}
\end{aligned}
$$

表 7　共通強度別強度差合計とその総和

共通強度別強度差合計	$(a_1+\ a_2+\ a_3+\cdots+\ a_{n-2}+\ a_{n-1}+\ a_n-\qquad na_1\quad)+\qquad(a_1\quad -a_1)$ $\{a_2+\ a_3+\cdots+\ a_{n-2}+\ a_{n-1}+\ a_n-(n-1)a_2\quad\}+\qquad(2a_2\quad -a_1-a_2)$ $\{a_3+\cdots+\ a_{n-2}+\ a_{n-1}+\ a_n-(n-2)a_3\quad\}+\qquad(3a_3\quad -a_1-a_2-a_3)$ $\vdots$ $\vdots$ $(a_{n-2}+\ a_{n-1}+\ a_n-\qquad 3a_{n-2})+\{(n-2)a_{n-2}-a_1-a_2-\cdots-a_{n-3}-a_{n-2}\}$ $(a_{n-1}+\ a_n-\qquad 2a_{n-1})+\{(n-1)a_{n-1}-a_1-a_2-\ \cdots\cdots\cdots\ -a_{n-2}\ -a_{n-1}\}$ $(a_n-\qquad a_n\quad)+\qquad(na_n\quad -a_1-a_2-\ \cdots\cdots\cdots\cdots\cdots\cdots\ -a_{n-1}-a_n)$
総和	$2\{na_n+(n-1)a_{n-1}+(n-2)a_{n-2}+\cdots+3a_3+2a_2+a_1-na_1-(n-1)a_2-(n-2)a_3-\cdots-3a_{n-2}-2a_{n-1}-a_n\}$

$$=2\{na_n+(n-1)a_{n-1}+(n-2)a_{n-2}+\cdots$$
$$+3a_3+2a_2+a_1$$
$$-na_1-(n-1)a_2-(n-2)a_3-\cdots$$
$$-3a_{n-2}-2a_{n-1}-a_n\} \tag{5}$$

以下では，ジーニに従って（p.21），強度の個数 n が①奇数のとき(T_O)と②偶数のとき(T_E)とに分けて，総和 T を計算する。

① n が奇数のとき

$$T_O=2\{na_n+(n-1)a_{n-1}+(n-2)a_{n-2}+\cdots$$
$$+\left(\frac{n+1}{2}+2\right)a_{\frac{n+1}{2}+2}+\left(\frac{n+1}{2}+1\right)a_{\frac{n+1}{2}+1}$$
$$+\left(\frac{n+1}{2}\right)a_{\frac{n+1}{2}}$$
$$+\left(\frac{n+1}{2}-1\right)a_{\frac{n+1}{2}-1}+\left(\frac{n+1}{2}-2\right)a_{\frac{n+1}{2}-2}$$
$$+\cdots+3a_3+2a_2+a_1$$
$$-na_1-(n-1)a_2-(n-2)a_3-\cdots$$
$$-\left(\frac{n+1}{2}+2\right)a_{\frac{n+1}{2}-2}-\left(\frac{n+1}{2}+1\right)a_{\frac{n+1}{2}-1}$$
$$-\left(\frac{n+1}{2}\right)a_{\frac{n+1}{2}}$$
$$-\left(\frac{n+1}{2}-1\right)a_{\frac{n+1}{2}+1}-\left(\frac{n+1}{2}-2\right)a_{\frac{n+1}{2}+2}$$
$$-\cdots-3a_{n-2}-2a_{n-1}-1a_n\}$$
$$=2[(n-1)a_n+\{(n-1)-2\}a_{n-1}$$
$$+\{(n-2)-3\}a_{n-2}+\cdots$$
$$+\left\{\left(\frac{n+1}{2}+2\right)-\left(\frac{n+1}{2}-2\right)\right\}a_{\frac{n+1}{2}+2}$$
$$+\left\{\left(\frac{n+1}{2}+1\right)-\left(\frac{n+1}{2}-1\right)\right\}a_{\frac{n+1}{2}+1}$$
$$+\left\{\left(\frac{n+1}{2}\right)-\left(\frac{n+1}{2}\right)\right\}a_{\frac{n+1}{2}}$$
$$-\left\{\left(\frac{n+1}{2}+1\right)-\left(\frac{n+1}{2}-1\right)\right\}a_{\frac{n+1}{2}-1}$$
$$-\left\{\left(\frac{n+1}{2}+2\right)-\left(\frac{n+1}{2}-2\right)\right\}a_{\frac{n+1}{2}-2}-\cdots$$
$$-\{(n-2)-3\}a_3-\{(n-1)-2\}a_2$$
$$-(n-1)a_1]$$
$$=2\{(n-1)a_n+(n-3)a_{n-1}+(n-5)a_{n-2}$$
$$+\cdots+4a_{\frac{n+1}{2}+2}+2a_{\frac{n+1}{2}+1}$$
$$+0\cdot a_{\frac{n+1}{2}}$$
$$-2a_{\frac{n+1}{2}-1}-4a_{\frac{n+1}{2}-2}$$
$$-\cdots-(n-5)a_3-(n-3)a_2-(n-1)a_1\}$$
$$=2\{(n-1)(a_n-a_1)$$
$$+(n-3)(a_{n-1}-a_2)$$
$$+(n-5)(a_{n-2}-a_3)$$
$$+\cdots$$
$$+4(a_{\frac{n+1}{2}+2}-a_{\frac{n+1}{2}-2})$$
$$+2(a_{\frac{n+1}{2}+1}-a_{\frac{n+1}{2}-1})$$
$$+0\cdot a_{\frac{n+1}{2}}\}$$
$$=2\left\{\sum_{i=1}^{\frac{n+1}{2}-1}(n+1-2i)(a_{n-i+1}-a_i)\right\}$$
$$+2\{0\times(a_{n-\frac{n+1}{2}+1}-a_{\frac{n+1}{2}})\}$$
$$=2\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i) \tag{5'}$$

② n が偶数のとき

$$T_E=2\{na_n+(n-1)a_{n-1}+(n-2)a_{n-2}+\cdots$$
$$+(\frac{n}{2}+2)a_{\frac{n}{2}+2}+(\frac{n}{2}+1)a_{\frac{n}{2}+1}$$
$$+(\frac{n}{2})a_{\frac{n}{2}}+(\frac{n}{2}-1)a_{\frac{n}{2}-1}+\cdots$$
$$+3a_3+2a_2+1a_1$$
$$-na_1-(n-1)a_2-(n-2)a_3-\cdots$$
$$-(\frac{n}{2}+2)a_{\frac{n}{2}-1}-(\frac{n}{2}+1)a_{\frac{n}{2}}$$
$$-(\frac{n}{2})a_{\frac{n}{2}+1}-(\frac{n}{2}-1)a_{\frac{n}{2}+2}-\cdots$$
$$-3a_{n-2}-2a_{n-1}-1a_n\}$$
$$=2\{(n-1)a_n+(n-3)a_{n-1}+(n-5)a_{n-2}$$
$$+\cdots+3a_{\frac{n}{2}+2}+1a_{\frac{n}{2}+1}$$

$$-1a_{\frac{n}{2}}-3a_{\frac{n}{2}-1}$$
$$-\cdots-(n-5)a_3-(n-3)a_2-(n-1)a_1\}$$
$$=2\{(n-1)(a_n-a_1)+(n-3)(a_{n-1}-a_2)$$
$$+(n-5)(a_{n-2}-a_3)+\cdots$$
$$+3(a_{\frac{n}{2}+2}-a_{\frac{n}{2}-1})+1(a_{\frac{n}{2}+1}-a_{\frac{n}{2}})\}$$
$$=2\sum_{i=1}^{\frac{n}{2}}(n+1-2i)(a_{n-i+1}-a_i) \qquad (5'')$$

(3)　強度差の個数

すでに述べたように，平均差 $\varDelta$ とは，強度差の総和 T を強度差の個数 N で除したときに得られる商のことである。したがって，強度差の相加平均としての平均差をもとめるには，(i) 強度差の総和 T と (ii) 強度差の個数 N の 2 つが必要である。(i) 強度差の総和 T については，強度の個数 n が奇数の場合［(5')式］と偶数の場合［(5'')式］に分けて，すでにもとめた。しかし，ここでは，強度の個数 n が奇数か偶数かは問題ではない。強度差の総和 T が(5)式[9] であたえられることだけを確認すれば，後は強度差の個数 N で強度差の総和 T を割ることによって平均差をもとめることができる。

そこで，(ii) 強度差の個数 N を考察する。強度差の総和 T のなかには，同一の強度どうしの強度差 a_i-a_i が算入されている（表 1～表 6 参照)。このために，どの共通強度 a_i についても強度差の個数は n である。そして，共通強度の個数は n である（表 1～表 6 参照)。このことから，強度差の総和 T は，$n\times n=n^2$ 個の強度差についての総和であることが分かる。本稿では，この T が $N=n^2$ 個の強度差の総和であることを明示する目的から，それを T_{n^2} で表す。

ところで，同一の強度どうしの強度差 a_i-a_i はつねにゼロである ($a_i-a_i=0$)。したがって，共通強度ごとの強度差の合計から，この強度差 a_i-a_i を取り除いても，すべての強度差の総和 T の値には変化はない。すなわち，上で見たように，T_{n^2}（同一の強度どうしの強度差を算入して得た強度差の総和）をもとめたときの強度差の個数 N は n^2 個であるが，この個数 n^2 から強度差がゼロになる強度差 (a_i-a_i) の個数を取り除き，残った強度差だけでその総和 T をもとめても，その値は変わらないのである。むしろ，そのときには，現実的意味に乏しい強度差 (a_i-a_i) を算入してもとめた総和 T_{n^2} よりも，実質的に意味のある強度差に限定してその総和を得ることができる。

同一の強度の組からなる強度差 (a_i-a_i) は，共通強度ごとに 1 個あって 1 個しかない。このような強度差を除去すれば，どの共通強度についても実質的に意味をもつ強度差の個数は $(n-1)$ 個になる。共通強度の個数が n 個，それぞれの共通強度について実質的な強度差が $(n-1)$ 個であるから，すべての共通強度にかんする実質的な強度差の個数 N は $n(n-1)$ 個になる $[N=n(n-1)]$（表 1～表 6 参照)。このときの強度差の総和 T が $n(n-1)$ 個の強度差であたえられることを明示する目的で，この総和をとくに $T_{n(n-1)}$ と書くことにする。

以上から，同一の強度差を算入するときの強度差の個数は $N=n^2$ となり，算入しない場合には $N=n(n-1)$ になる。同一の強度どうしの強度差はゼロであるから，これを算入した n^2 個の強度差の総和 T_{n^2} と算入しない $n(n-1)$ 個の強度差の総和 $T_{n(n-1)}$ は等しい。すなわち

$$T_{n^2}=T_{n(n-1)}=T \qquad (6)$$

である。

後述するように，ジーニは，強度差の個数

9)　T
$$=2\{na_n+(n-1)a_{n-1}+(n-2)a_{n-2}+\cdots$$
$$+3a_3+2a_2+a_1$$
$$-na_1-(n-1)a_2-(n-2)a_3-\cdots$$
$$-3a_{n-2}-2a_{n-1}-a_n\} \qquad (5)$$

N を n^2 とする場合だけでなく，$n(n-1)$ とする場合についても平均差の計算式を導出している。しかし，集中比（ジーニ係数）を再定義するときに用いた平均差 $\varDelta$ では，強度差の個数が $n(n-1)$ となっている。その理由は必ずしも明確ではない。ここでは，同一の強度どうしの強度差に実質的意味を見出しにくいために，ジーニはそれを除外したと解釈したい。そして，集中比との関連で用いられる平均差については，その個数が $n(n-1)$ であることをここで改めて確認しておく。

2．さまざまな平均差

(1)　平均差

ジーニは強度差の総和 T を，強度の個数 n が奇数の場合［(5')式］と偶数の場合［(5")式］に分けて誘導した。このために，強度差の総和 T を強度差の個数 N［$=n(n-1)$］で除して得られる商（強度差の相加平均）としての平均差を一般に $\varDelta$ で表すと，$\varDelta$ は，n が奇数のときと偶数のときのそれぞれについてあたえられることになる。

強度の個数 n が奇数の場合の平均差 $\varDelta_O$ は，強度差の総和 T_O［(5')式］を強度差の個数 N［$=n(n-1)$］で除した

$$\varDelta_O=\frac{T_O}{n(n-1)}=\frac{2}{n(n-1)}\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i)$$

によってあたえられる。

他方で，n が偶数のときには，平均差 $\varDelta_E$ は，強度差の総和 T_E［(5")式］を強度差の個数 N［$=n(n-1)$］で除した

$$\varDelta_E=\frac{T_E}{n(n-1)}=\frac{2}{n(n-1)}\sum_{i=1}^{\frac{n}{2}}(n+1-2i)(a_{n-i+1}-a_i)$$

によってあたえられる。

強度の個数 n が偶数か奇数かを問わなければ，強度差の相加平均 $\varDelta$（すなわち平均差）は，一般に，

$$\varDelta=\frac{2}{n(n-1)}\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i) \quad (4)［再掲］$$

であたえられ，(4)式が誘導されたことになる（p. 22）。(4)式において n が偶数の場合には，最終項を規定する $\frac{n+1}{2}$ は $\left(\frac{n}{2}+\frac{1}{2}\right)$ となり，整数 $\frac{n}{2}$ よりも 0.5 大きい値をあたえるが，端数 0.5 が付く順位の項は存在しないので $\frac{n+1}{2}$ は実際には $\frac{n}{2}$ となって，（偶数の場合の）$\varDelta_E$ は（奇数の場合の）$\varDelta_O$ と同値（$\varDelta$）になるからである。

なお，ここで，後の考察のために，強度の個数 n が奇数か偶数を問わないときの，強度差の総和 T をもとめておく。強度差は全部で $n(n-1)$ 個あるので，強度差の平均としての平均差 $\varDelta$［(4)式］を $n(n-1)$ 倍すれば，相加平均の補償機能によって強度差の総和 T を得ることができる。すなわち，

$$T=2\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i) \quad (7)$$

である。これが

$$\begin{aligned}T=2\{&na_n+(n-1)a_{n-1}+(n-2)a_{n-2}+\cdots\\&+3a_3+2a_2+a_1\\&-na_1-(n-1)a_2-(n-2)a_3-\cdots\\&-3a_{n-2}-2a_{n-1}-a_n\}\end{aligned} \quad (5)［再掲］$$

と同値であることは言うまでもない。

表 8　共通強度別完全非重複強度差一覧表

共通強度	a_1	a_2	a_3	……	a_{n-2}	a_{n-1}
強 度 差	a_2-a_1 a_3-a_1 a_4-a_1 ・ ・ a_n-a_1	a_3-a_2 a_4-a_2 ・ ・ a_n-a_2	a_4-a_3 ・ ・ a_n-a_3	……	$a_{n-1}-a_{n-2}$ a_n-a_{n-2}	a_n-a_{n-1}
強度差の個数	$n-1$	$n-2$	$n-3$	……	2	1

⑵　完全非重複平均差

共通強度別の表 1 ～表 6（強度差とその合計，前出）を見れば，次のことが分かる。

(a)　強度差のなかには，同一の強度どうしの強度差 a_i-a_i が 1 個ずつあること。

(b)　共通強度を a_1 とする強度差（表 1 ）のなかには，それ以外の強度差を共通強度とする表（表 2 ～表 6 ）に重複して存在しているものがあること。

たとえば，共通強度を a_2 とする表 2 には，表 1 に記載されている a_2-a_1 がある（重複的強度差の個数は 1 個）。また，共通強度 a_3 の表（表 3 ）には，表 1（共通強度 a_1）の強度差 a_3-a_1 と表 2（共通強度 a_2）の強度差 a_3-a_2 が重複して記載されている（重複的強度差の個数は 2 個）。以下，重複的強度差の個数は 1 個ずつ増えてゆく。

そこで，(i) 強度差とその合計にかんするすべての表（表 1 ～表 6 参照）から共通強度どうしの強度差を削除し，さらに，(ii) 表 2 以降の表からは，共通強度を a_1 とする強度差にかんする表 1 にも重複して記載されている強度差を削除する。そして，共通強度別に強度差を表にまとめることにする（表 8 ）。

この表 8 に一覧される強度差には，(i) 共通強度どうしの強度差と (ii) 一度でも計数された強度差とが記載されていないという二重の意味で，重複がない。このために，表 8 に記載された強度差を「完全非重複強度差」と言うことにする。

表 8 から明らかなように，完全非重複強度差の個数の和は，

$$(n-1)+(n-2)+(n-3)+\cdots\cdots+2+1$$

であり，

$$\{(n-1)+1)\times(n-1)\times\frac{1}{2}=\frac{n(n-1)}{2}$$

となる。すなわち，表 8 における強度差の個数は，ジーニが強度差の総和 T［(6)式参照］をもとめたときの 1/2 である[10)]。

表 8 においては，総和 T をもとめたときと較べて強度差の個数が半分になっているということは，表 8 における強度差の総和を T_P とおけば，この T_P も T の半分であることを意味する。すなわち，

$$T_P=\frac{1}{2}T \tag{8}$$

であり，したがって，

$$T=2T_P \tag{9}$$

10）強度差に類似した概念と考えられる観測差［ 2 つの観測値の差］の個数をこのように $\frac{n(n-1)}{2}$ と規定したのは，W. ヨルダン（1869 年）であり，この観測差の総和にかんする一般式をあたえたのは，C.G. フォン・アンドレ（1872 年）である。この点については次節（ 3 ．平均差概念の形成）で述べる。

である。

平均差は強度差の相加平均であるから，完全非重複強度差にかんする平均差を $\varDelta_P$ で表すと，

$$\varDelta_P=\frac{T_P}{\dfrac{n(n-1)}{2}}$$

$$=\frac{\dfrac{1}{2}T}{\dfrac{n(n-1)}{2}} \qquad \text{[(8)式による]}$$

$$=\frac{T}{n(n-1)}=\varDelta \tag{10}$$

となる。

同一の強度どうしの強度差を除外し，かつまた，一度でも強度差の総和に算入された強度差は，これを除外するという，二重の意味で重複を回避したときの平均差 $\varDelta_P$ を「完全非重複平均差」（differenza media perfetta senza ripetizione）と名づけることにする。このとき，$\varDelta_P$ はジーニがもとめた平均差 $\varDelta$ と同値であることを⑽式は示している。

なお，平均差 $\varDelta$ にあっては同一の強度どうしの強度差が強度差の総和 T に算入されていないので，「完全非重複平均差」$\varDelta_P$ と識別するために，$\varDelta$（ジーニの平均差）を「非重複平均差」（differenza media senza ripetizione）と言えば，事柄が明確になる。

⑶　重複平均差

$\varDelta$ [⑷式[11]]（非重複平均差）をもとめるとき，その前段で強度差の総和 T をもとめた。その総和 T には同一の強度どうしの強度差 a_i-a_i が合算されている。$a_i-a_i=0$ であり，共通強度ごとにこのような強度差は1個ずつあるから，共通強度ごとに見られる同一の強度どうしの強度差の個数を，共通強度別の強度差の個数から減ずれば，実質的な強度差は共通強度ごとに $(n-1)$ 個あることになる。共通強度は全部で n 個あるので，同一の強度どうしの強度差を含まない実質的な強度差の個数は $n(n-1)$ 個となる。この個数 $n(n-1)$ を除数，強度差の総和 T を被除数として得られる商が（非重複）平均差 $\varDelta$ であることは上に述べたとおりである。

これにたいして，同一の項どうしの強度差も強度差の個数に算入することにすれば，すでに述べたように，強度差の個数 N は，$n^2\,(=n\times n)$ である。ところが，同一の強度どうしの強度差は，これもまたすでに述べたように，いずれもゼロであり，これを強度差の総和 T に算入しても，T の値は⑺式と同じ値となる。このとき，平均差 $\varDelta_R$ は，⑺式があたえる強度差の総和 T を強度差の個数 n^2 で割ることによって，次のように得ることができる。

$$\varDelta_R=\frac{T}{n^2}$$

$$=\frac{2}{n^2}\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i) \tag{4'}$$

この平均差 $\varDelta_R$ のことをジーニは「n 個の諸量の重複平均差（differenza media con ripetizione tra le n quantitȧ）」と言っている（p.22）。

平均差 $\varDelta$（非重複平均差）をあたえる⑷式と重複平均差 $\varDelta_R$ をあたえる(4')式から，2種類の平均差の比の値は

$$\frac{\varDelta_R}{\varDelta}=\frac{\dfrac{2}{n^2}\displaystyle\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i)}{\dfrac{2}{n(n-1)}\displaystyle\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i)}$$

$$=\frac{n-1}{n}$$

となる。

11）$\varDelta=\dfrac{2}{n(n-1)}\displaystyle\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i)$ 　(4)

したがって，同一の強度どうしの強度差をその個数には計数しない非重複平均差 $\varDelta$［(4)式］と計数する重複平均差 $\varDelta_R$［(4’)式］との関係は

$$\varDelta_R = \frac{n-1}{n}\varDelta \tag{11}$$

または

$$\varDelta = \frac{n}{n-1}\varDelta_R \tag{12}$$

である（p.22)。(11)式を変形して，

$$\varDelta_R = \left(1-\frac{1}{n}\right)\varDelta \tag{13}$$

とすれば，$n \to \infty$ のとき，

$$\varDelta_R = \varDelta \tag{14}$$

となることが分かる。すなわち，十分大きな n については非重複平均差 $\varDelta$ と重複平均差 $\varDelta_R$ は同値である。

エマヌエル・ツーバーはジーニにならって平均差を，「平均差（die mittlere Differenz)」$\varDelta$ と「重複平均差（die mittlere Differenz mit Wiederholung)」$\varDelta'$［ジーニの $\varDelta_R$］との 2 つに分類し，両者の間に(11)式が成立することを述べた後，「私にしてみれば，［$\varDelta'$よりも］$\varDelta$［非重複平均差］の方が理にかなっている（naturgemäß）ように見える。……私としては今後とも $\varDelta$ を使用するつもりである。」と述べている[12)]。

12）Czuber, Emanuel, “Beitrag zur Theorie statistischer Reihen,” *Versicherungswissenschaftlichen Mitteilungen*, Vol.9, 1914, p.122.

(4)　さまざまな平均差の数学的関係（簡単な要約）

以上の考察から，強度差の総和とそれをもとめるために算入した強度差の個数に応じて，平均差は次のように分類できる（ただし，n は強度の個数)。

平均差
- 非重複平均差 $\varDelta$
  ……ジーニが集中比の定義に用いた平均差（ジーニの平均差）
  ［強度差の個数は $n(n-1)$］
- 完全非重複平均差 $\varDelta_P$
  ……二重の意味で重複する強度差を除外した平均差
  ［強度差の個数は $\frac{1}{2}n(n-1)$］
- 重複平均差 $\varDelta_R$
  ……ジーニの重複平均差
  ［強度差の個数は n^2］

そして，それぞれの平均差の関係は次のようになる。

$$\begin{cases} \varDelta = \varDelta_P & (10)［再掲］\\ \varDelta = \dfrac{n}{n-1}\varDelta_R & (12)［再掲］\\ \varDelta = \varDelta_R \ （n\text{が十分大きいとき}） & (14)［再掲］\end{cases}$$

よって，強度の個数 n が十分に大きい場合には，

$$\varDelta = \varDelta_P = \varDelta_R \tag{15}$$

が成立する。

ここで，平均差が強度差の相加平均であるという定義に立ち返り，そして，強度差を絶対値 $|a_i - a_j|$ で表すことにする。そうすると，ジーニが平均差（非重複平均差）$\varDelta$ をもとめたときの強度差の総和 T は今日の表記法では，次のように表すことができる。

$$T=\sum_{i<j}2\times|a_i-a_j|$$
$$=2\sum_{i<j}|a_i-a_j| \qquad (16)$$

ただし，$i=1,2,\cdots\cdots,n-1;$
$j=i+1,i+2,\cdots\cdots,n$

同様に，完全非重複平均差 $\varDelta_P$ については，強度差の総和 T_P が

$$T_P=\frac{1}{2}T \qquad (8)\text{［再掲］}$$
$$=\sum_{i<j}|a_i-a_j| \qquad \text{［(16)式による］}(17)$$

ただし，i，j については(16)式に同じ。

となる。

また，重複平均差 $\varDelta_R$ については，強度差の総和 T_R は

$$T_R=\sum_{i=1}^{n}\sum_{j=1}^{n}|a_i-a_j| \qquad (18)$$

である。平均差をもとめるときに計数した強度差の個数で，それぞれの総和 T，T_P，T_R を割ると，三種類の平均差は次のようになる。

①　非重複平均差

$$\varDelta=\frac{T}{n(n-1)}$$
$$=\frac{\sum_{i<j}2\times|a_i-a_j|}{n(n-1)}$$
$$=\frac{2\sum_{i<j}|a_i-a_j|}{n(n-1)} \qquad (19)$$

②　完全非重複平均差

$$\varDelta_P=\frac{T_P}{\frac{n(n-1)}{2}}$$
$$=\frac{\frac{1}{2}T}{\frac{n(n-1)}{2}} \qquad \text{［(8)式による］}$$
$$=\frac{T}{n(n-1)}$$
$$=\varDelta$$
$$=\frac{2\sum_{i<j}|a_i-a_j|}{n(n-1)} \qquad \text{［(19)式による］}$$
$$=\frac{\sum_{i<j}|a_i-a_j|}{\frac{n(n-1)}{2}} \qquad (20)$$

③　重複平均差

$$\varDelta_R=\frac{T_R}{n^2}$$
$$=\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_i-a_j|}{n^2} \qquad (21)$$

一般に $\varDelta_P=\varDelta$［(10)式］が成立し，とくに強度の個数 n が十分大きいときは，$\varDelta_R=\varDelta$［(14)式］が成立するので，

$$\varDelta=\varDelta_P=\varDelta_R \qquad (15)\text{［再掲］}$$

となる。このため，(19)式，(20)式，(21)式のどれを用いて平均差を計算しても，違いがないように考えられる。(19)式はジーニがあたえる平均差の公式を今日の表記法で記述し直したものであるが，平均差の概念に実質的意味をもって整合するのは(20)式（完全非重複平均差 $\varDelta_P$）である。その理由は，同一の値をあたえる強度差を重複して計算することに積極的意義を見出しにくく，強度差の計算にあっては重複の回避が望ましいということにある。しかも，同一の強度どうしの組から得られる強度差はゼロであり，そこには強度差としての実質的な意義はない。このことから，ジーニの見解に反することにはなるが，平均差を計算するときには，同一の強度どうしの強度差を除外して強度差の個数を計数することが要請されていると考えられる。重複する強度差を除外することは，ガウス没後の 19 世紀誤差論でも一般的に見られる（ただし，後述

するように，そこでは観測差と言われている)。その意味では，⒇式があたえる完全非重複平均差 Δ_P は誤差論における理論展開に整合する。

そこで，次に項を改めて，19 世紀誤差論における「観測差」や「平均差」をめぐる論議を検討する。そして，ジーニ理論と誤差論との理論的関連を検討する手がかりを得たい。

3．平均差概念の形成

⑴　誤差論と平均差

強度を観測値に置き換えて，平均差の概念を捉え直すと，平均差は観測値間の差の相加平均と見なすことができる。同一の物理的対象にかんする観測値を 2 つずつ組み合わせて，その差をとり，その相加平均をもとめるという数学的操作だけに着目すれば，平均差は，19 世紀中葉以降の観測値結合論（いわゆる誤差論）のなかにも見出すことができる。1876 年に公刊された F.R. ヘルメルトの論文[13]（以下 1876 a 年論文と略記）では，W. ヨルダン[14]によって提起された観測精度の確定問題が取り上げられている。これを解く過程でヘルメルトは正規分布に従う 2 つの観測誤差の差にかんする相加平均を用いているからである。後述するように，2 つの観測誤差の差は 2 つの観測値の差（観測差）に等しい。その限りでは，たとえヘルメルトが観測誤差の差の相加平均を単に「平均値」と言っていようとも，その「平均値」はジーニの平均差に対応しているかに見える。このためであろうか，平均差の概念はジーニの独創ではなくて，ガウス以降の 19 世紀誤差論ですでに研究されていたとの指摘がある[15]。H.A. デイビッドは次のように述べている[16]。

> …… g をジーニの平均差（Gini's Mean Difference）と呼ぶことにするが，本質的にはこれと同一の統計量は 1876 年にヘルメルトによってすでに論じられていた。しかし，それは当時としてもけっして真新しいものではなかった！ヘルメルトが数式の普遍妥当性を確定したのは事実であるが，その もとい をなしたのは，彼よりもわずかばかり先に『天文通報』（*Astronomische Nachrichten*）誌上で意見を交換していた W. ヨルダンとフォン・アンドレである。

このことを裏づけるかのように，ジーニが平均差を論じた 1912 年論文ではヘルメルトの異なる 2 篇の論文がそれぞれ別の箇所で引用されている（p.19，p.59)。そのうち最初に引用されているヘルメルト論文は『数学・物理学雑誌』（第 21 巻）に掲載された「観測誤差の累乗和の確率ならびに関連問題について」[17]（以下これを 1876 b 年論文と略記）である。この論文の刊行年は 1876 年であるが，本節冒頭において平均差との関連で掲げたヘルメルトの 1876 a 年論文（「等精度の直接観測値の確率誤差の計算におけるペーテルスの公式の厳密性」）とは別の論文である。ヘルメルトの 1876 b 年論文は，正規分布に従う

13）Helmert, Friedrich Robert, "Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Fehlers directer Beobachtungen gleicher Genauigkeit," *Astronomische Nachrichten*, Vol.88, 1876. ［Helmert (1876a)］

14）Jordan (1869).

15）① David, H.A., "Gini's Mean Difference rediscovered," *Biometrika*, Vol.55, 1968, p.573f. ［David (1968)］; ② ditto, "Early Sample Measurement of Variability," *Statistical Science*, Vol.13, No.4, 1998, p.378f. ［David (1998)］

16）David (1968), p.573f.

17）Helmert, F.R., "Ueber die Wahrscheinlichkeit der Potenzsummen der Beobachtungsfehler und über einige damit in Zusammenhange stehende Fragen," *Zeitschrift für Mathematik und Physik*, Vol.21, 1876.

分布の分散（標準偏差）を主題としている[18]。それを考察することによって初めて標本分散の分布が明らかになった[19]。

理論史上はそのように位置づけられているヘルメルトの 1876 b 年論文を，ジーニはその 1912 年論文において

（i）正規分布に従う観測値の誤差 e の分布の尺度

$$^{m}S_{A}=\sqrt[m]{\frac{\sum|e_i|^m}{n}} \qquad (3)［再掲］$$

が観測値の個数 n の増大に伴って限界値 L に近づくこと，

（ii）(3)式が L に近づく速度（rapidità）は，$m=2$ のときに最大になること，

を指摘する文脈で引用している（p.19）。彼は，ヘルメルトの 1876 b 年論文をいわゆる第 1 カテゴリーに属す変動性の尺度との関連で参照しているのである。

他方，平均差（第 2 カテゴリーに属す変動性の尺度）をめぐる論議に一応の決着をつけたとしてデイビッドが掲げている[20]（ヘルメルトの）1876 a 年論文は，ジーニの 1912 年論文の別の箇所（p.59）で引用されている（これについては本節(6)で述べる）。

ところが，最近になって，デイビッドは，19 世紀（とりわけガウス以降の）誤差論とジーニの平均差との関連に言及し，先行研究とは独立にジーニが平均差を再発見したという趣旨の発言をしている[21]。このデイビッドの見解の検討を含めて，ジーニの平均差理論と 19 世紀誤差論における理論展開との間のさらに詳しい関連については今後の研究課題である。さしあたり，この点にかんしてはウムベルト・リッチの次の見解[22]が，おおむね妥当すると思われる。

> 平均差は一部の数学者によってすでに知られていたが，経済統計の論議，すなわち所得分布の研究への応用を提起した勲功はフルランに帰する。しかしながら，彼は手短に示唆しただけである。幅広く，体系的な論述はジーニによってなされた。

基本的にはそのとおりであろう。しかし，ジーニによる平均差の誘導と 19 世紀誤差論者の研究との間の共通点や類似点，そして異同を検討することによって，断片的であるにせよ，両者の理論的関連を窺い知ることができる。そこで以下では，両者の見解を対比することだけに課題を絞り込んで，ガウス以降の 19 世紀誤差論における理論的展開を限定的に取り上げ，誤差論者とジーニとの理論的紐帯の一端を考察してみたい。ここで限定した論点とは，（i）観測差と強度差，（ii）観測差の個数と強度差の個数，（iii）観測差の総和と強度差の総和，（iv）観測差の相加平均と強度差の相加平均の 4 点である。論点が散漫になることを防ぐ意味からも，誤差論の細部にわたる数理的展開には触れることを避けたいからに他ならない。なお，結論的には先に引用したリッチの見解を覆すものではないことをあらかじめ断っておく。

(2) W. ヨルダン(1)（1869 年）

誤差論の分野では，（i）観測装置の性能の向上，（ii）装置の操作方法にたいする習熟，（iii）観測条件の管理の 3 条件をクリアしてもなおかつ，個々の観測値は微妙に異なるこ

18）David (1998), p.371f.

19）小河原正巳「標本分布論」中山伊知郎編『統計学辞典（増補版）』東洋経済新報社　1957 年　p.215。

20）David (1998), p.374.

21）David (1998), p.374.

22）Ricci, Umberto, “L’Indice di variabilità e la curva dei redditi,” *Giornale degli Economisti e Rivista di Statistica*, Vol.53, 1916, p.192. ［Ricci (1916)］

とが問題とされる。客観的に真値が存在するとすれば，観測値(v)は真値(t)と観測誤差(ε)の合成であると考えられる（$v=t+\varepsilon$)。正規分布に従う観測誤差の分布の分散 σ^2 を 2 倍して，その逆数 $\frac{1}{2\sigma^2}$ を h^2 とおいたとき，この h を精度指標と言う[23]。したがって，h で計測される観測精度とは観測値の再現性の尺度であり，これが大きいほど（σ^2 が小さいほど）似たような値の観測値が得られることになる。ガウスの最小二乗法では，この精度指標 h の値が未知であっても，観測値の相加平均はもとめるべき真値の推定値と見なされる[24]。

カールスルーエ大学の教授であった W. ヨルダンによれば，そのガウスは「誤差を最小にする観測値の組み合わせ理論」の第 2 部(1823 年)[25] のなかで次のように述べていると言われている。

> この方法［観測精度の近似的確定方法］は，十分に多数の真の観測誤差が厳密に既知であるということを前提としている。この条件は厳格に言えば，けっしてとは言わないまでも，ほとんど満たされることがない。観測の精度評価に事後的（ア・ポステリオリ）にかかわろうとして計算するすべての人々は，最小二乗法によってもとめられる未知量の値を真値と見なす方法に従っている。だが，これは明らかに理論的には間違っていて，多くの場合には，たとえ実際的な目的にかなっているとしても，重大な誤りを犯すことがある。したがって，この問題の厳密な取り扱いには最大限の努力があってしかるべきである（強調はヨルダン)。

ガウスのこの問題提起を受けたヨルダンは，「精度の厳密な確定」をみずからの課題として，ガウスの死（1855 年）の 14 年後に「単一の未知量にかんする多数回反復観測の精度の確定について」という論文を執筆した。それは『天文通報』誌（通巻 1766-67 号，1869 年）に掲載された[26]。以後，観測精度の厳密な確定をめぐる論議がこの誌上で行われ，それとの関連で 1876 年のヘルメルト論文（1876 a 年論文）では精度の尺度として活用されてきた（「中央誤差（mittlerer Fehler)」とも言われる）「確率誤差（der wahrscheinliche Beobachtungsfehler)」ρ（$-\rho$ から $+\rho$ までの誤差分布の確率が 1/2 となる値で，$N(0,1)$ のときには $\rho=0.6745$）が検討され，ヨルダンの問題提起にたいして一応の結着がついた。その後，精度の確定問題は統計的推定（推定量の数学的特性）の問題と重なり，20 世紀に入っても論議されているので[27]，1876 a 年論文によるヘルメルトの解答をもって最終段階に到達したとは言い難いが，ρ をもとめる過程でジーニの平均差に似た概念が形成された（これについては後述する)。

さて，論争の口火を切った 1869 年論文のなかで，ヨルダンは，同一の物理的対象にたいする「2 つの観測値の差 d の確率（die Wahrscheinlichkeit der *Differenz d zwier* Beobachtungen)」を考察した。そして，この d の個数が全部で $\frac{n(n-1)}{2}$ であると述

23）Hogben, L., *Statistical Theory*, New York 1957, p.164. ［Hogben (1957)］(木村和範訳『統計の理論』梓出版社　1986 年　p.164。)

24）Hogben (1957), p.189f.［訳書 p.187f.］

25）Gauss, C. F., "Theoria combinationis observationum erroribus minimis obnoxiae," (1823). ただし，引用は Jordan (1869), p.216 のドイツ語訳による。『カール・F. ガウス　誤差論』(飛田武幸・石川耕春訳）紀伊國屋書店　1981 年　p.48 参照。

26）Jordan (1869).

27）宇野利雄「誤差論」中山伊知郎編『統計学辞典（増補版)』東洋経済新報社　1957 年　p.475 以下，とくに p.477 以下。

べ[28)]，この差 d を「観測差（Beobachtungs-differenz）」と名づけた[29)]。

この概念を手がかりにして，ヨルダンは「精度の厳密な確定」を志向したが，1869 年論文ではそれは果たされていない。ここでは，ヨルダンの観測差を強度差と言い換えれば，用語としては，強度差の語源はそこにあると言いうることに注意したい。観測差の相加平均を計算するには（i）観測差の総和と（ii）観測差の個数が必要である。この 2 つのうちのひとつ［(ii)］がヨルダンによってあたえられたのである。また，ヨルダンとジーニでは取り扱う数量的規定性が異なっていることにも注意したい。すなわち，ヨルダンの取り扱った観測値は正規分布に従うのにたいして，ジーニの強度は分布の正規性を前提としない。誤差論で取り扱う数量の特殊性を確認した上で，観測差の確率や観測精度を確定することの必要性を主張したのは，C.G. フォン・アンドレであった。

⑶　C.G. フォン・アンドレ⑴（1869 年）

ヨルダンの論文が掲載された『天文通報』（通巻 1766-67 号）を読んだ翌日（1869 年 8 月 11 日），デンマーク王国の枢密顧問官フォン・アンドレは同誌の編集者に宛てて，ヨルダン論文を批判する書簡を送った。デンマーク語で書かれたその書簡は同誌通巻 1770 号に掲載された[30)]。その批判の要点は，観測値（および観測誤差）の正規性をヨルダンが考慮していないということであった。

この書簡のなかで，フォン・アンドレは，ヨルダンが全部で $\dfrac{n(n-1)}{2}$ 個あるとした観測差を「偶発誤差（virkelige Feil）$\varDelta$ の差」と言い換えて，$\varDelta$ の差を下のような「三角表（triangulære Tableau）」にまとめた[31)]。これによって観測差の個数にかんするヨルダンの見解がより明証的になり，後にヨルダンもこの三角表を使用するようになった。ただし，平均差をジーニに従ってすでに $\varDelta$ と表記したので，本稿では混同を避ける目的でフォン・アンドレの $\varDelta$ を ε で表すことにする。

$$
\begin{array}{lllll}
\varepsilon_2-\varepsilon_1 & & & & \\
\varepsilon_3-\varepsilon_1 & \varepsilon_3-\varepsilon_2 & & & \\
\varepsilon_4-\varepsilon_1 & \varepsilon_4-\varepsilon_2 & \varepsilon_4-\varepsilon_3 & & \\
\varepsilon_5-\varepsilon_1 & \varepsilon_5-\varepsilon_2 & \varepsilon_5-\varepsilon_3 & \varepsilon_5-\varepsilon_4 & \\
\vdots & \vdots & \vdots & \vdots & \\
\varepsilon_n-\varepsilon_1 & \varepsilon_n-\varepsilon_2 & \varepsilon_n-\varepsilon_3 & \varepsilon_n-\varepsilon_4\cdots\cdots & \varepsilon_n-\varepsilon_{n-1}
\end{array}
$$

この三角表から明らかなように，フォン・アンドレは（ヨルダンとともに），同一の誤差どうしの差 $(\varepsilon_i-\varepsilon_i)$ を除外している。さらにまた，偶発誤差を昇順に並べることによって，小さくない誤差 ε_j と大きくない誤差 ε_i との乖離 $(\varepsilon_j-\varepsilon_i)$ だけを考察の対象としていることも分かる。この点では，共通強度ごとに可能な N 個の強度差を枚挙し，それらの総和 T をもとめたジーニの手法とは異なっている（1.⑵強度差の総和参照）。フォン・アンドレにあっては観測誤差の差の個数は，ジーニの強度差の個数 $N\,[=n(n-1)]$ の半分になっている。このために，観測誤差の差の総和もジーニの強度差の総和 T の 1/2 である（$T/2$）。

上述のように，観測誤差の差を強度差に擬制することができるのは，次の理由による。本稿における表記上の統一を図るために，観

28）Jordan (1869), p.216. ただし，原文では n が m と表記されている。

29）Jordan (1869), p.217.

30）von Andrae, C.G., “Schreiben des Herrn Geheimen Etatsraths *von Andrä* an den Herausgeber,” *Astronomische Nachrichten*, Vol.74, 1869. ［von Andrae (1869)］

31）von Andrae (1869), p.284. ただし，原文では n が m と表記されている。

測値を v，真値を t，偶発誤差を ε とおくと，任意の 2 つの観測値は次のようになる。

$$v_i = t + \varepsilon_i$$
$$v_j = t + \varepsilon_j$$
ただし，$v_i \leqq v_j$

したがって，観測差 $v_j - v_i$ は

$$\begin{aligned} v_j - v_i &= (t+\varepsilon_j) - (t+\varepsilon_i) \\ &= \varepsilon_j - \varepsilon_i \end{aligned} \tag{22}$$

となる。単一の真値 t が存在していることを前提すれば，観測差 $(v_j - v_i)$ は偶発誤差の差 $(\varepsilon_j - \varepsilon_i)$ と等しい。このために，フォン・アンドレの三角表に記載された偶発誤差の差は，ヨルダンがもとめようとした観測差と同一である。この観測差を強度差に置き換えれば，強度差はこの観測差に匹敵する。この限りでは，強度差はヨルダン（1869 年）の観測差と同一であるだけでなく，フォン・アンドレ（1869 年）の観測誤差の差とも同一である。ただし，すでに述べたように，総和をもとめるべき差の個数には違いがある。その個数をヨルダンやフォン・アンドレは $\dfrac{n(n-1)}{2}$ としたが，ジーニは $n(n-1)$ とした。

この段階では，ヨルダンが提起した「観測精度の確定」問題にたいする解答はもとより，平均差を計算するために必要となる観測差の総和もあたえられていない。ただし，観測差の個数にかんするヨルダンの見解が三角表によってより明確になったという点では，フォン・アンドレはヨルダンの所説を一歩先に進めたと言える。

(4) W. ヨルダン(2)（1872 年）

フォン・アンドレは 1869 年書簡で観測値の正規性を閑却しているとヨルダンを批判したことはすでに指摘した。それにたいして，ヨルダンは「観測値の誤差にふさわしい法則に従う観測値が対をなすときに作る差を取り扱うことは，事情によっては利点がある」（強調はヨルダン）ことを主張したいのであって，けっして観測値が正規分布に従うことを無視してはいないと反論した[32]。

そして，彼は，偶発誤差の差にかんするフォン・アンドレの三角表に記載される被減数と減数のサフィックスを入れ替えた。後に示すようにヨルダンの目的が差の平方和をもとめることであったために，サフィックスの入れ替えは問題がないと考えたのかもしれない。あるいは，誤差を降順に並べて，（フォン・アンドレと同様に）小さくない誤差を被減数，大きくない誤差を減数とおいたのかもしれない。ヨルダンもまたフォン・アンドレと同様に偶発誤差を $\varDelta$ とおいたが，表記を統一するために，ここでもそれを ε で表した三角表を下に掲げる。

$$\begin{array}{lll}
\varepsilon_1 - \varepsilon_2 & & \\
\varepsilon_1 - \varepsilon_3 & \varepsilon_2 - \varepsilon_3 & \\
\varepsilon_1 - \varepsilon_4 & \varepsilon_2 - \varepsilon_4 & \\
\quad\vdots & \quad\vdots & \\
\varepsilon_1 - \varepsilon_n & \varepsilon_2 - \varepsilon_n \cdots\cdots\cdots & \varepsilon_{n-1} - \varepsilon_n
\end{array}$$

ヨルダンはこの $(\varepsilon_i - \varepsilon_j)$（ただし，$i<j$）の二乗和を $[dd]$ と表記した。二乗和 $[dd]$ をもとめるために，上の三角表から

$$\begin{array}{lll}
(\varepsilon_1 - \varepsilon_2)^2 & & \\
(\varepsilon_1 - \varepsilon_3)^2 & (\varepsilon_2 - \varepsilon_3)^2 & \\
(\varepsilon_1 - \varepsilon_4)^2 & (\varepsilon_2 - \varepsilon_4)^2 & \\
\quad\vdots & \quad\vdots & \\
(\varepsilon_1 - \varepsilon_n)^2 & (\varepsilon_2 - \varepsilon_n)^2 \cdots\cdots\cdots & (\varepsilon_{n-1} - \varepsilon_n)^2
\end{array}$$

32) Jordan, W., "Ueber die Bestimmung des mittleren Fehlers durch Wiederholung der Beobachtungen," *Astronomische Nachrichten*, Bd. 79, 1872, p.219. ［Jordan (1872)］

という新しい三角表を作成する。そのかっこを解くと，次のような三角表が得られる。

$$
\begin{array}{lll}
\varepsilon_1^2-2\varepsilon_1\varepsilon_2+\varepsilon_2^2 & & \\
\varepsilon_1^2-2\varepsilon_1\varepsilon_3+\varepsilon_3^2 & \varepsilon_2^2-2\varepsilon_2\varepsilon_3+\varepsilon_3^2 & \\
\varepsilon_1^2-2\varepsilon_1\varepsilon_4+\varepsilon_4^2 & \varepsilon_2^2-2\varepsilon_2\varepsilon_4+\varepsilon_4^2 & \\
\vdots & \vdots & \\
\varepsilon_1^2-2\varepsilon_1\varepsilon_n+\varepsilon_n^2 & \varepsilon_2^2-2\varepsilon_2\varepsilon_n+\varepsilon_n^2 \cdots\cdots & \varepsilon_{n-1}^2-2\varepsilon_{n-1}\varepsilon_n+\varepsilon_n^2
\end{array}
$$

この三角表を用いると $(\varepsilon_i-\varepsilon_j)$ の二乗和 $[dd]$ は次のようになる。

$$
\begin{aligned}
[dd]=&(n-1)\varepsilon_1^2+(\varepsilon_2^2+\varepsilon_3^2+\varepsilon_4^2+\cdots+\varepsilon_n^2)\\
&-2(\varepsilon_1\varepsilon_2+\varepsilon_1\varepsilon_3+\varepsilon_1\varepsilon_4+\cdots+\varepsilon_1\varepsilon_n)\\
&+(n-2)\varepsilon_2^2+(\varepsilon_3^2+\varepsilon_4^2+\cdots+\varepsilon_n^2)\\
&-2(\varepsilon_2\varepsilon_3+\varepsilon_2\varepsilon_4+\cdots+\varepsilon_2\varepsilon_n)\\
&+\cdots\cdots\\
&+\varepsilon_{n-1}^2+\varepsilon_n^2\\
&-2\varepsilon_{n-1}\varepsilon_n\\
=&(n-1)(\varepsilon_1^2+\varepsilon_2^2+\varepsilon_3^2+\varepsilon_4^2+\cdots+\varepsilon_n^2)\\
&-2\sum_{i=2}^{n}\varepsilon_1\varepsilon_i-2\sum_{i=3}^{n}\varepsilon_2\varepsilon_i-\cdots\cdots-2\sum_{i=n}^{n}\varepsilon_{n-1}\varepsilon_i
\end{aligned}
\tag{23}
$$

ヨルダンは「これらの値の二乗和では（周知の前提により）積の２倍は相殺しあう。」と述べ，もとめるべき二乗和 $[dd]$ を

$$
[dd]=(n-1)(\varepsilon_1^2+\varepsilon_2^2+\varepsilon_3^2+\varepsilon_4^2\cdots+\varepsilon_n^2) \tag{24}
$$

とした[33]。そして，この二乗和 $[dd]$ を差の個数 $\frac{n(n-1)}{2}$ で割り，その商の平方根を D とおいた。すなわち，D は

$$
D=\sqrt{\frac{[dd]}{\frac{1}{2}n(n-1)}} \tag{25}
$$

である。ヨルダンはこの D を「２つの観測値の平均差（*mittlere Differenz zweier Beobachtungen*）」（強調はヨルダン）と言っている[34]。

このヨルダンの「平均差」の数理的意味を明確にする目的で，(25)式に(24)式を代入して整理する。

$$
\begin{aligned}
D&=\sqrt{\frac{[dd]}{\frac{1}{2}n(n-1)}}\\
&=\sqrt{\frac{(n-1)(\varepsilon_1^2+\varepsilon_2^2+\varepsilon_3^2+\varepsilon_4^2+\cdots+\varepsilon_n^2)}{\frac{1}{2}n(n-1)}}\\
&=\sqrt{\frac{\sum_{i=1}^{n}\varepsilon_i^2}{\frac{1}{2}n}}\\
&=\sqrt{\frac{2\sum_{i=1}^{n}\varepsilon_i^2}{n}}\\
&=\sqrt{2}\sqrt{\frac{\sum_{i=1}^{n}\varepsilon_i^2}{n}}
\end{aligned}
\tag{26}
$$

ヨルダンは，(26)式を誘導してはいない。しかし，(26)式によれば，ヨルダンの「平均差」D は，観測誤差 ε の平方 ε^2 にかんする相加平均 $\frac{\sum_{i=1}^{n}\varepsilon_i^2}{n}$（平方平均）の平方根の $\sqrt{2}$ 倍になっていることが分かる。これは明らかにジーニが導出した平均差（われわれのいわゆる非重複平均差）Δ とは異なり，さらにまた，完全非重複平均差 Δ_P や重複平均差 Δ_R とも異なっている。それだけでなく，ヨルダンの「平均差」D と観測精度の確定（h の算定）との関係については議論の余地を残す統計量であることも分かる。ヨルダンが措定した課題（精度の確定）は未解決のままであった。確かに，「平均差」はジーニに先立つこと40

33）Jordan (1872), p.219.

34）Jordan (1872), p.219.

年前に提起された。しかし，ヨルダンとジーニとではそれぞれ別の統計量を指していることに注意したい。

⑸ C.G. フォン・アンドレ⑵（1872 年）

ヨルダンの上記論文（1872 年）を読んだフォン・アンドレは，『天文通報』の編集者に宛てた先の 1869 年書簡における主張を展開する必要から，同誌に論文[35]を投稿した（1872 年）。この論文では，観測誤差の差の総和とともに観測値そのものの差（ヨルダンの観測差）の総和があたえられている。この総和を観測差の個数で割れば，観測差の相加平均が得られる。強度と観測値の概念を数量的規定性一般にまで抽象するとき，この相加平均は平均差と見なしうる。観測差の総和があたえられれば，これをヨルダンがもとめた観測差の個数で割ることによって平均差が得られるので，平均差の概念へは後一歩の所まで到達したと言うことができる。

フォン・アンドレは n 個の観測値を a_i とおいた。そして，そこから 2 つの観測値で組を作り，そのうち小さくない観測値から大きくない観測値を引いたときの差を，観測誤差にかんする三角表（上述）と同様の表にまとめた。ただし，本稿ではジーニの強度を a で表したので，混同を避けるために，ここではアンドレの観測値 a_i を v_i で表すことにする（原文では n が m と書かれている）。

$$
\begin{array}{lllll}
v_2-v_1 & & & & \\
v_3-v_1 & v_3-v_2 & & & \\
v_4-v_1 & v_4-v_2 & v_4-v_3 & & \\
v_5-v_1 & v_5-v_2 & v_5-v_3 & v_5-v_4 & \\
\vdots & \vdots & \vdots & \vdots & \\
v_n-v_1 & v_n-v_2 & v_n-v_3 & v_n-v_4 \cdots\cdots & v_n-v_{n-1}
\end{array}
$$

⑵式[36]により，観測差の総和は観測誤差の差の総和に等しい。すなわち

$$\textstyle\sum_{i<j}(v_j-v_i)=\sum_{i<j}(\varepsilon_j-\varepsilon_i) \qquad (22')$$

である。フォン・アンドレはこの総和を $[d]$ とおいた。そして，観測差にかんする $[d]$ と観測誤差の差にかんする $[d]$ を次のようにもとめた。

観測差の和は

$$\begin{aligned}[d]=&(n-1)\{v_n-v_1\}+(n-3)\{v_{n-1}-v_2\}\\ &+(n-5)\{v_{n-2}-v_3\}+\cdots\cdots \qquad (27)\end{aligned}$$

である。また，観測誤差の差の総和については(22')式から同様に，

$$\begin{aligned}[d]=&(n-1)\{\varepsilon_n-\varepsilon_1\}+(n-3)\{\varepsilon_{n-1}-\varepsilon_2\}\\ &+(n-5)\{\varepsilon_{n-2}-\varepsilon_3\}+\cdots\cdots \qquad (28)\end{aligned}$$

を得る。

ヘルメルトに継承された計算式は㉗式およびとくに㉘式である。ジーニが対象としたのは強度差であるから，平均差（ジーニ）との関連では観測差にかんする㉗式が注目される。そこで，フォン・アンドレの㉗式を今日の表記法で表せば，

$$[d]=\sum(n+1-2i)(v_{n-i+1}-v_i) \qquad (27')$$

35) von Andrae, C.G., “Ueber die Bestimmung des wahrscheinlichen Fehlers durch die gegebenen Differenzen von m gleich genauen Beobachtungen einer Unbekannten,” *Astronomische Nachrichten*, Vol.79, p.259f. [von Andrae (1872)]

36) $v_i=t+\varepsilon_i$（ただし，v_i は観測値，t は真値，ε_i は観測誤差）であるから，一般に

$$\begin{aligned}v_j-v_i&=(t+\varepsilon_j)-(t+\varepsilon_i)\\ &=\varepsilon_j-\varepsilon_i \qquad (22)\end{aligned}$$

となる。

となる。フォン・アンドレの三角表から明らかに，観測差の個数は $\frac{n(n-1)}{2}$ である。すでに述べたように，ジーニは強度差の総和 T として

$$T=2\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i) \qquad (7)\,[\text{再掲}]$$

を誘導した。そのときの強度差の個数 N は $n(n-1)$ であった。したがって，観測差と強度差がいずれも２つの数量的規定性の差を示しているという共通性に着目すれば，フォン・アンドレの $[d]$［(27')式］とジーニの T［(7)式］の間の関係は，

$$T=2\times[d]$$

であることが分かる。

要するに，強度差を基準強度別に悉皆的に（ただし，同一強度間の強度差を除く）枚挙して $n(n-1)$ 個の強度差の総和をもとめるか（ジーニ），あるいは一度でも観測差を合計のなかに算入させた場合にはそれを観測差の総和の計算から除外して $\frac{n(n-1)}{2}$ 個の観測差についてその総和をもとめるか（ヨルダン，フォン・アンドレ）が，両者の違いである。

以上の指摘はジーニ理論から見た誤差論についてであるが，誤差論の分野に限定すれば，フォン・アンドレによっても，ヨルダンが提起した精度の確定問題への解答はまだあたえられてはいない。

(6)　F.R. ヘルメルト（1876 年）

①「平均誤差」

アーヘン（ドイツ）のポリテヒニクム（高等工業学校）教授であったヘルメルトは，フォン・アンドレの結果を一部活用して，ヨルダンの提起した問題（観測精度の確定）にたいする解答を試みた[37]。

ヘルメルトは『天文通報』通巻 2096-97 合併号に掲載された 1876 a 年論文のなかで，ヨルダンとフォン・アンドレの研究の帰結を次のように要約している。

(a)　$\frac{n(n-1)}{2}$ 個の可能な観測差（Beobachtungsdifferenz）に着目して，その合計がもとめられていること，

(b)　可能な観測差を d とすると，その和 $[d]$ は

$$[d]=(l_n-l_1)(n-1)+(l_{n-1}-l_2)(n-3)+(l_{n-2}-l_3)(n-5)+\cdots\cdots$$

ただし，l_i は観測値（$i=1, 2, \cdots\cdots, n$）

であること［(27)式参照］，

(c)　$[d]$ をあたえる上式はフォン・アンドレによること。

(d)　この式を用いれば，「よりたやすく（bequemer）」観測差の総和を計算できること。

以上のような先行研究を踏まえたヘルメルトの解答（精度の確定）については，エマヌエル・ツーバー[38]による分かりやすい解説

37）Helmert (1876a), p.128. なお David (1998), p. 374 によれば，以下の叙述の典拠は Helmert (1876a)であるとされている。しかし，Helmert (1876a)には，①“Die Berechnung des wahrscheinlichen Beobachtungsfehlers aus den Quadraten der Verbesserungen directer Beobachtungen gleicher Genauigkeit und die Fechnersche Formel”［pp.120ff.］と②“Die Berechnung des wahrscheinlichen Beobachtungsfehlers aus den ersten Potenzen der Differenzen gleichgenauer directer Beobachtungen”［pp.127ff.］の２つの論文が付帯されている。これらは本文と関連した論点を取り上げてはいるが，それとしては独立の論文とも見なしうる。以下の議論は②でなされている。

38）Czuber, Emanuel, *Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung*, Leib-

がある。そこで以下ではそれに従い，ヘルメルトの所説を敷衍する。ただし，今日の一般的な表記法を用いる。

ツーバーは，真値を X（未知），観測値を l，偶発誤差を ε，観測差を $\varDelta$ とおいている。しかし，本稿では表記上の統一を図るために，ツーバーと同様に偶発誤差は ε で示すが，X を t で，l を v で，$\varDelta$ を d で表す。このときには，

$$\begin{aligned} t &= v_1+\varepsilon_1 = v_2+\varepsilon_2 \\ d &= v_1-v_2 = \varepsilon_2-\varepsilon_1 \end{aligned} \tag{29}$$

となる[39]。

ツーバーによれば，差 d の分布は

$$\frac{h}{\sqrt{2\pi}}\,e^{-\frac{h^2}{2}d^2}$$

ただし，h は精度指標，π は円周率，e は自然対数の底。

に従うので，dの絶対値 $|d|$（これを abs. d と表記する）の分布の平均を $\overline{\text{abs.}\,d}$ とおくと

$$\overline{\text{abs.}\,d}=\frac{1}{h\sqrt{\pi}}\sqrt{2} \tag{30}$$

となる。

ここで，

$$\frac{1}{h\sqrt{\pi}}=\vartheta \tag{31}$$

とおく。そして，(31)式を(30)式に代入すると，$|d|$ の平均 $\overline{\text{abs.}\,d}$ は

$$\overline{\text{abs.}\,d}=\vartheta\sqrt{2} \tag{32}$$

となる。

ところで，観測差 d が s 個$(d_1, d_2, \cdots\cdots, d_s)$ ある場合には，$|d|$ の平均 $\overline{\text{abs.}\,d}$ は，

$$\overline{\text{abs.}\,d}=\frac{\sum_{i=1}^{s}|d_i|}{s} \tag{33}$$

である。

観測差の個数 s は 2 つの観測値の対の個数に等しい。また，観測値が n 個のときには，その観測差の個数 s は

$$s=\frac{n(n-1)}{2} \tag{34}$$

である。

(34)式を(33)式に代入すれば，

$$\overline{\text{abs.}\,d}=\frac{\sum_{i=1}^{s}|d_i|}{\frac{n(n-1)}{2}} \tag{35}$$

を得る。

(35)式を(32)式に代入すれば，

$$\vartheta\sqrt{2}=\frac{\sum_{i=1}^{s}|d_i|}{\frac{n(n-1)}{2}}$$

となる。これを整理すれば，

$$\begin{aligned} \vartheta &= \frac{\sum_{i=1}^{s}|d_i|}{\frac{n(n-1)}{2}\sqrt{2}} \\ &= \frac{2\sum_{i=1}^{s}|d_i|}{n(n-1)\sqrt{2}} \\ &= \frac{2\sqrt{2}\sum_{i=1}^{s}|d_i|}{n(n-1)\sqrt{2}\sqrt{2}} \end{aligned}$$

となって，

zig 1903, p.233f. ［Czuber (1903)］

39) (29)式からも偶発誤差 ε の差は観測差に等しく，同一の値 (d) で表現できることが分かる。(22)式参照。

$$\vartheta=\frac{(\sum_{i=1}^{s}|d_i|)\sqrt{2}}{n(n-1)} \tag{36}$$

を得る。

以上がツーバーによる説明の骨子である。ヘルメルトもまた，ツーバーとは異なる数式展開によってではあるが，ツーバーと同様の⑶⑹式を誘導した[40]。そして，ヘルメルトは⑶⑹式があたえる統計量を「平均誤差（Durchschnittsfehler）ϑ」と名づけた。ツーバーによる説明は⑶⑹式ならびにそれと関連する数式の提示で終わっている。

ヘルメルトは 1876 a 年論文で⑶⑹式を導出した後，独自の数式展開から確率誤差（中央誤差）ρ をもとめ，精度の確定問題への解答を試みた[41]。しかし，ここではその解答（ρ）に立ち入らない。ヘルメルトの ϑ とジーニの $\varDelta$ との数学的関係を明らかにするだけで本稿の所期の目的を達成できるからである。その鍵をあたえるのが⑶⑸式と⑶⑹式である[42]。そこで項を改めて，そのことについて述べる。

② ϑ（ヘルメルトの平均誤差）と $\varDelta$（ジーニの平均差［非重複平均差]）

ヘルメルトの ϑ とジーニの $\varDelta$ との間の数学的関係について考察するために，

$$\overline{\text{abs.}\,d}=\frac{\sum_{i=1}^{s}|d_i|}{\frac{n(n-1)}{2}} \tag{35 [再掲]}$$

に着目する。⑶⑸式の分子（観測差の絶対値の総和）は，(i) 同一の強度どうしの強度差を

---

40) Helmert (1876a), p.129.

41) Helmert (1876a), pp.129ff.

42) ⑶⑹式を手がかりにすれば，ヘルメルトとは別の仕方で精度の確定問題にたいする解答を見出すことができる。そこで，⑶⑴式を次のように変形する。

$$h=\frac{1}{\vartheta\sqrt{\pi}} \tag{31'}$$

この(31')式に⑶⑹式でもとめた平均誤差 ϑ を代入すれば，精度指標 h を得ることができる。すなわち，h は

$$h=\frac{1}{\frac{(\sum_{i=1}^{s}|d_i|)\sqrt{2}}{n(n-1)}\sqrt{\pi}}$$

$$=\frac{1}{\frac{(\sum_{i=1}^{s}|d_i|)\sqrt{2}\cdot\frac{1}{2}}{n(n-1)\cdot\frac{1}{2}}\sqrt{\pi}}$$

$$=\frac{1}{\frac{\sum_{i=1}^{s}|d_i|}{\frac{n(n-1)}{2}}\cdot\frac{\sqrt{2\pi}}{2}}$$

$$=\frac{1}{\frac{\sum_{i=1}^{s}|d_i|}{\frac{n(n-1)}{2}}\cdot\sqrt{\frac{\pi}{2}}}$$

$$=\frac{1}{\frac{\sum_{i=1}^{s}|d_i|}{\frac{n(n-1)}{2}}}\cdot\sqrt{\frac{2}{\pi}} \tag{*}$$

によってあたえられる。(*)式右辺の第 1 項は観測差の相加平均 $\overline{\text{abs.}\,d}$ ［⑶⑸式，$\overline{\text{abs.}\,d}=\frac{\sum_{i=1}^{s}|d_i|}{\frac{n(n-1)}{2}}$ (再掲)］の逆数であるから，(*)式によって精度指標 h は $\overline{\text{abs.}\,d}$ の逆数を $\sqrt{\frac{2}{\pi}}$ 倍すればあたえられることが分かる。ヘルメルトは確率誤差 ρ によって精度を確定しようと考えたので，これは彼の解答ではないが，観測差にもとづく観測精度の確定というヨルダンの問題提起にたいするひとつの解答と考えることができる。

あるいは，(*)式の逆数をとって，

$$\frac{1}{h}=\frac{\sum_{i=1}^{s}|d_i|}{\frac{n(n-1)}{2}}\cdot\sqrt{\frac{\pi}{2}}$$

とすれば，精度指標 h の逆数が観測差の相加平均 $\overline{\text{abs.}\,d}$ の $\sqrt{\frac{\pi}{2}}$ 倍になっていると言うこともできる。

排除し，(ii) 2 つの強度を重複なく組み合わせ，(iii) 小さくない強度から大きくない強度を引いたときに計算される「完全非重複強度差」の総和 T_P に等しい。また，分母は「完全非重複強度差」の個数に等しい。したがって，(35)式は「完全非重複平均差」$\varDelta_P$ に等しい。すなわち，

$$\overline{\mathrm{abs.}\,d}=\varDelta_P$$

である。しかも，

$$\varDelta=\varDelta_P \qquad (10)［再掲］$$

により，

$$\overline{\mathrm{abs.}\,d}=\varDelta \qquad (37)$$

であるから，統計量 $\overline{\mathrm{abs.}\,d}$ は，ジーニの平均差 $\varDelta$（非重複平均差）に等しい。

ところが，上で見たように，

$$\vartheta=\frac{(\sum_{i=1}^{s}|d_i|)\sqrt{2}}{n(n-1)} \qquad (36)［再掲］$$

である。ここで，(35)式を変形すれば，

$$\overline{\mathrm{abs.}\,d}=\frac{2\sum_{i=1}^{s}|d_i|}{n(n-1)} \qquad (35')$$

$$\therefore\quad \frac{\overline{\mathrm{abs.}\,d}}{2}=\frac{\sum_{i=1}^{s}|d_i|}{n(n-1)} \qquad (35'')$$

となる。そして，(35″)式を(36)式に代入すれば

$$\vartheta=\frac{\overline{\mathrm{abs.}\,d}}{2}\sqrt{2}$$
$$=\frac{1}{\sqrt{2}}\overline{\mathrm{abs.}\,d} \qquad (38)$$

となる。さらに，(38)式に(37)式を代入すれば，

$$\vartheta=\frac{1}{\sqrt{2}}\varDelta \qquad (39)$$
$$\varDelta=\sqrt{2}\vartheta \qquad (39')$$

となる。すなわち，ジーニの平均差 $\varDelta$ はヘルメルトの平均誤差 ϑ の $\sqrt{2}$ 倍になっている。

なお，正規分布に従う観測値にかんしては，平均偏差（lo scostamento semplice medio）${}^1S_A$ の $\sqrt{2}$ 倍と平均差 $\varDelta$ が等しくなること（$\varDelta=\sqrt{2}\,{}^1S_A$）について「厳密な証明をあたえた」論者としてジーニはヘルメルトの名を挙げている（p.59）。1912 年論文のなかで，ジーニはヘルメルトの論文を 2 篇引用していることはすでに述べたが，第 2 論文（1876 a 年論文の付帯論文[43]）はこの文脈で引用されている。統計量 ${}^1S_A$ を ϑ と読み替えるジーニの叙述は簡潔で難解である。この点の検討は今後の課題である。ここではヘルメルトの立論が諸量の正規性を前提していることに，ジーニはその特質を見ていると指摘するにとどめる。

(7) V. フルラン（1911 年）

リッチは，上述した研究者の他に平均差概念に関連のある業績を残したとして V. フルラン[44] の名を挙げている。誤差論（観測精

43) 脚注37の②（Helmert, F. R., “Die Berechnung des wahrscheinlichen Beobachtungsfehlers aus den ersten Potenzen der Differenzen gleichgenauer Beobachtungen”）参照。

44) Furlan, V., “Neue Literatur zur Einkommensverteilung in Italien,” *Jahrbücher für Nationalökonomie und Statistik*, III. Folge, 42. Band, 1911, p.255. ［Furlan (1911)］なお，フルランについては次も参照。① von Bortkiewicz, Ladislaus, “Die Disparitätsmasse der Einkommensstatistik,” *Bulletin of the International Statistical Institute,* Vol. 25, no.3, 1931, p.198；②汐見三郎「分配均等度の測定」中山伊知郎編『統計学辞典（増補版）』東洋経済新報社　1957 年　p.646。

度の確定問題）を研究したと言うよりは，むしろ所得分布の統計的計測にかんする業績を残したフルランは，所得 x 以上の人数を $f(x)$ とする「所得曲線（Einkommenskurve）」を

$$y=f(x) \tag{40}$$

と定義している。そして，2つの所得 x と x' $(x<x')$ にかんする二重積分の絶対値

$$\left|2\int_u^\infty\int_x^\infty (x'-x)\,\frac{dy}{dx}\,\frac{dy'}{dx'}\,dx\,dx'\right| \tag{41}$$

ただし，dy と dy' の絶対値はそれぞれ所得が x から $x+dx$ までの人数と，所得が x' から $x'+dx$ までの人数。また，u は捕捉した最低所得。

が「所得分布における不平等性の指数（Index）ないし測度（Maß）」であるとフルランは述べている。所得についてその差 $(x'-x)$ をとり，それをもって所得分布の不平等度を計測する試みに限定すれば，このフルランの論文は，平均差を定式化したジーニの 1912 年論文の前年に公刊されているので，ジーニに直接先行していると言うことができる。しかし，フルランは平均差というよりは，むしろそれをもとめるときに，その前段で必要とされる所得差（強度差）とその総和に着目したのであって［(41)式参照］，しかもその指摘が示唆にとどまることは上に見られるとおりである。したがって，この問題について「幅広く，体系的」（リッチ）に論述したのはジーニであるとのリッチの見解は妥当なものと考えることができる。

(8)　ジーニの平均差理論と誤差論

本節の最後に用語と数式に分けて，ジーニの平均差概念と誤差論との理論的関連に言及する。

①　用語

(a)　観測差と強度差

ヨルダンは，「観測差（Beobachtungsdifferenz）」によって観測精度を確定しようと試みた。物理的対象を観測する目的は，それに固有の一意的な真値をもとめることにある。そして，観測値には正規分布に従う偶発誤差が内在することを前提として，誤差論は理論的展開を見た。真値の存在を前提する誤差論では，観測差は偶発誤差の差に等しい。ここで，観測差（もしくは偶発誤差の差）を，自然と社会を問わず広く客観的対象の数量的規定性にかんする2つの数値の差というレベルにまで一般化すると，その限りではヨルダンの観測差はジーニの「n 個の諸量の差（differenza tra n quantità）」（本稿に言う強度差）と異ならない。ジーニ理論と誤差論は，いずれも2つの数量的規定性の差にかんする統計量を考察したという意味では，ジーニの平均差の基本的な構想は，ヨルダンからフォン・アンドレを経てヘルメルトにいたる誤差論の延長線上に位置づけることができる。ジーニはこれらの誤差論研究者が「かなり以前から，まったく異なった観点で多数の諸量間の平均差を研究してきた」ことを認めている（p.58）。ただし，彼らが研究対象とした諸量は「ガウス曲線（la curva di Gauss）」に従っており，そのような諸量を前提にして平均差を誘導する過程ではたとえ似たような数式が導出されることがあろうとも，1912 年論文で展開した所説はヨルダンからヘルメルトにいたる誤差論研究とは「独立していて，しかもまったく異なる」とジーニは主張している（p.59）。

(b)　平均差

ジーニは $n(n-1)$ 個の強度差の相加平均を平均差（differenza media）と名づけ，Δ で表した。また，n^2 個の強度差の相加平均を重複平均差（differenza media con ripetizione）と名づけ，$\varDelta_R$ で表した。

これにたいして，ヨルダンの平均差（mittlere Differenz）は，観測誤差の平方平均の平方根を $\sqrt{2}$ 倍した値を意味し，言葉としては同一ではあるが，いかなる意味でもジーニの平均差とは異なっている。

② 数式

(a) 強度差の個数と観測差の個数

ジーニの場合，平均差をもとめるべき強度差の個数は，$\varDelta$（非重複平均差）のときには $n(n-1)$ 個であり，$\varDelta_R$（重複平均差）のときには n^2 個である。$\varDelta$ の計算には同一の強度どうしの強度差が除去されているが，$\varDelta_R$ の場合にはそれが除去されない。

これにたいして，ヨルダンからヘルメルトにいたる誤差論の分野では，観測差の総和に算入される観測差のなかには，同一の観測値どうしの観測差はもとより，一度でも計算された観測差が重複して算入されることはない。そのために，観測差の個数はジーニの平均差 $\varDelta$ の場合の半分となって，$\frac{n(n-1)}{2}$ である。

(b) 強度差の総和と観測差の総和

強度差と観測差を 2 つの数量的規定性の差のレベルにまで抽象して，そこに強度差と観測差との共通性を見出すとすれば，強度差の総和を T とおくとき，観測差の総和は $T/2$ と表すことができる。上で述べたように，観測差の総和を計算するときの観測差の個数は，強度差の総和をもとめるときの 1/2 だからである。

ジーニが強度差の総和をもとめるときに誘導した一般式と同様の数式は，すでにフォン・アンドレが誘導している。ジーニはその先行研究の存在を知っていた。しかし，フォン・アンドレの取り扱う諸量が正規分布に従っている点や証明の仕方の点で違いがあると述べ，ジーニはみずからの独創性を主張している（p.59)。

(c) 強度差の相加平均と観測差の相加平均

(b)で示したそれぞれの総和を，(a)で述べたそれぞれの差の個数で割れば，それぞれの相加平均が得られる。

強度差の相加平均は，

$$\frac{T}{n(n-1)} \quad \text{または} \quad \frac{T}{n^2}$$

ただし，n が十分に大きいときは $n-1 \rightarrow n$ となって，$\frac{T}{n(n-1)}=\frac{T}{n^2}$

である。

また，観測差の相加平均は，

$$\frac{\frac{T}{2}}{\frac{n(n-1)}{2}}=\frac{T}{n(n-1)}$$

である。

したがって，十分大きな n については，強度差の相加平均は観測差の相加平均と一致する。

ジーニにとっては，強度差の相加平均をもとめることが目的であり，彼はそれを平均差と名づけたが，ヨルダン（観測差の個数の確定）からフォン・アンドレ（観測差の総和の確定）を経てヘルメルト（観測差の相加平均の導出・援用）にいたる誤差論研究者にとっては，上述の相加平均をもとめることが最終目的ではなく，観測精度の確定途上で得られた「副産物」にすぎない。

③ 小括

本節における考察をヘルメルトとジーニの違いに限定すれば，次のように要約することができる。第 1 に，ヘルメルトは，偶発誤差と偶発誤差との乖離の研究に重点を置いた。彼が取り上げた誤差 ε とは「真の誤差

（wahre Fehler）」であり，観測値のなかにひそむとされる（いわゆる誤差論［観測値結合論］における）偶発誤差であり，そこでは真値の存在が前提されている。これにたいして，ジーニは強度に誤差を前提することなく，強度そのものを言わば真値として分析の対象とした。

第2に，ジーニは，いわゆる強度が正規分布に従うという前提をおいていない。これにたいして，ヘルメルト（をはじめとする誤差論研究者）では，観測値分布の正規性が重要な役割を果たしている。

上述の2点がジーニとヘルメルトを分かつ相違点となっている。ただし，このこととともに，ジーニ理論と（ヘルメルトを含む）19世紀誤差論との理論的関連については次の点を指摘しておく必要もある。ヘルメルトは対をなす偶発誤差の乖離について，その総和と個数から相加平均をもとめ，それを「平均値（Duchschnittswerth）」と言っている。この「平均値」は平均差に対応するので，ヘルメルトは平均差という言葉を使ってはいないが，彼もまたも平均差をもとめた（あるいは，ジーニ以前に平均差をもとめた）と言うことができる。

したがって，強度と強度の乖離について（相加）平均（平均差）をもとめるという考え方，ならびにそれを計算するときのために必要な強度差の概念は（誤差論の分野では観測差という名で呼ばれていたが），19世紀中葉にはすでに形成されていた。この意味では平均差という構想そのものは，ジーニの創始になるものではなく，19世紀中葉における観測値結合論（いわゆる誤差論）の分野で真値（true value）にたいする推定の精度を測定・向上させるための研究のなかで形成された概念である。このように，ジーニ理論と誤差論とでは，強度差の総和 T と強度差の相加平均 $\varDelta$ の計算，および観測差の総和 $\sum_{i=1}^{s}|d_i|$ と観測差の相加平均 $\overline{\mathrm{abs.}\,d}$ の計算までは共通性を見出すことができる。

しかし，ジーニ理論では強度間格差の計測が問題とされたので，平均差 $\varDelta$ の計算は「終着駅」であった。これにたいして，ヘルメルトが $\varDelta=\sqrt{2}\vartheta$ の関係にある平均誤差 ϑ をもとめたのは，確率誤差（中央誤差）によって精度を確定するためであった。ジーニの意味での平均差 $\varDelta$ は誤差論者にしてみれば，最終目的地にいたるための「通過駅」である。

4．平均差によるジーニ係数の再定義

(1)　再定義のための準備

n 個の個体の強度 a_i の総和 A_n は

$$A_n=\sum_{i=1}^{n}a_i$$

（一般に，第 k 項までの強度の和を A_k とすれば，$A_k=\sum_{i=1}^{k}a_i$ となる。）

であり，その相加平均 M_n をあたえる式

$$M_n=\frac{A_n}{n}$$

から

$$2M_n=\frac{2A_n}{n} \tag{42}$$

を得る。集中比 R（ジーニ係数）を初めて取り扱った1914年論文によれば，このとき，集中比 R は，

$$R=\frac{\varDelta}{2M_n} \tag{43}$$

ただし，$\varDelta$ は(4)式で定義される平均差。

である[45]。すなわち，R はすべての強度の相加平均の2倍と平均差との比率としてあた

45）Gini (1914), p.1237.

えられる。以下では，1914 年論文におけるこの定義を敷衍する。

(43)式の証明の方針は以下のとおりである。まず，(43)式に(4)式[46]と(42)式を代入して，

$$R=\frac{\frac{2}{n(n-1)}\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i)}{\frac{2A_n}{n}}$$

$$=\frac{\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i)}{(n-1)A_n} \quad (44)$$

を得る。

(43)式と(44)式は同値であるから，(44)式の証明をもって，(43)式の証明とみなすことができる。ジーニは，(43)式の証明のために集中比 R の定義式[47]

$$R=\frac{\sum_{i=1}^{n-1}(p_i-q_i)}{\sum_{i=1}^{n-1}p_i}$$

から誘導される

$$R=1-\frac{2}{(n-1)A_n}\sum_{i=1}^{n-1}A_i \quad (45)$$

を用いた[48]。

ここで，(45)式を変形すれば

$$R=\frac{(n-1)A_n-2\sum_{i=1}^{n-1}A_i}{(n-1)A_n} \quad (45')$$

となるので，(44)式による集中比 R の定義を証明するには，(44)式と(45')式の分子に着目して，

(44)式の分子＝(45')式の分子

を証明すればよい。すなわち

$$\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i)$$
$$=(n-1)A_n-2\sum_{i=1}^{n-1}A_i \quad (46)$$

が成り立つことを証明すればよい。

(2)　平均差と集中比

上述の方針に沿って(44)式を証明するために，(46)式の右辺を整理する。

①　(46)式右辺の第 1 項 $[(n-1)A_n]$

$$\begin{aligned}
&(n-1)A_n\\
&=(n-1)(a_1+a_2+a_3+\cdots+a_n)\\
&=(n-1)a_1+(n-1)a_2+(n-1)a_3+\cdots\\
&\quad+(n-1)a_{n-2}+(n-1)a_{n-1}+(n-1)a_n\\
&\qquad+0\times(a_1-a_1)+1\times(a_2-a_2)+2\times(a_3-a_3)+\cdots\\
&\qquad+(n-3)(a_{n-2}-a_{n-2})+(n-2)(a_{n-1}-a_{n-1})\\
&\qquad+(n-1)(a_n-a_n)\\
&=(n-1)a_1+(n-2)a_2+(n-3)a_3\\
&\quad+\cdots+2\times a_{n-2}+1\times a_{n-1}+0\times a_n\\
&\qquad+a_2+2a_3\\
&\qquad+\cdots+(n-3)a_{n-2}+(n-2)a_{n-1}+(n-1)a_n
\end{aligned} \quad (47)$$

②　(46)式右辺の第 2 項 $[2\sum_{i=1}^{n-1}A_i]$

$$\begin{aligned}
&2\sum_{i=1}^{n-1}A_i\\
&=2(A_1+A_2+A_3+\cdots A_{n-1})\\
&=2(a_1\\
&\quad+a_1+a_2\\
&\quad+a_1+a_2+a_3\\
&\qquad\vdots\\
&\quad+a_1+a_2+a_3+\cdots\cdots+a_{n-1})\\
&=2\{(n-1)a_1+(n-2)a_2+\cdots\cdots+2a_{n-2}+a_{n-1}\}
\end{aligned} \quad (48)$$

46）$\mathit{\Delta}=\frac{2}{n(n-1)}\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i)$　(4)

47）Gini (1914), p.1207.

48）Gini (1914), p.1238.

③　(46)式の右辺 $[(n-1)A_n-2\sum_{i=1}^{n-1}A_i]$

(47)式と(48)式を(46)式の右辺に代入すれば，次式を得る。

$$
\begin{aligned}
&(n-1)A_n-2\sum_{i=1}^{n-1}A_i\\
=&\{(n-1)a_1+(n-2)a_2+(n-3)a_3\\
&+\cdots+2\times a_{n-2}+1\times a_{n-1}+0\times a_n\\
&+a_2+2a_3\\
&+\cdots+(n-3)a_{n-2}+(n-2)a_{n-1}+(n-1)a_n\}\\
&-2\{(n-1)a_1+(n-2)a_2+(n-3)a_3\\
&+\cdots+2a_{n-2}+a_{n-1}\}\\
=&\{a_2+2a_3\\
&+\cdots+(n-3)a_{n-2}+(n-2)a_{n-1}+(n-1)a_n\}\\
&+\{(n-1)a_1+(n-2)a_2+(n-3)a_3\\
&+\cdots+2\times a_{n-2}+1\times a_{n-1}\}\\
&-2\{(n-1)a_1+(n-2)a_2+(n-3)a_3\\
&+\cdots+2\times a_{n-2}+1\times a_{n-1}\}\\
=&\{0\times a_1+1\times a_2+2\times a_3\\
&+\cdots+(n-3)a_{n-2}+(n-2)a_{n-1}+(n-1)a_n\}\\
&-\{(n-1)a_1+(n-2)a_2+(n-3)a_3\\
&+\cdots+2\times a_{n-2}+1\times a_{n-1}+0\times a_n\}\\
=&[\{na_n+(n-1)a_{n-1}+(n-2)a_{n-2}\\
&+\cdots+3\times a_3+2\times a_2+1\times a_1\}\\
&-(a_n+a_{n-1}+a_{n-2}+\cdots+a_3+a_2+a_1)]\\
&-\{(n-1)a_1+(n-2)a_2+(n-3)a_3\\
&+\cdots+2\times a_{n-2}+1\times a_{n-1}+0\times a_n\}\\
=&\{na_n+(n-1)a_{n-1}+(n-2)a_{n-2}\\
&+\cdots+3\times a_3+2\times a_2+1\times a_1\}\\
&-(a_n+a_{n-1}+a_{n-2}+\cdots+a_3+a_2+a_1)\\
&-\{(n-1)a_1+(n-2)a_2+(n-3)a_3\\
&+\cdots+2\times a_{n-2}+1\times a_{n-1}+0\times a_n\}\\
=&\{na_n+(n-1)a_{n-1}+(n-2)a_{n-2}\\
&+\cdots+3\times a_3+2\times a_2+1\times a_1\}\\
&-\{na_1+(n-1)a_2+(n-2)a_3\\
&+\cdots+3\times a_{n-2}+2\times a_{n-1}+1\times a_n\}\\
=&\{na_n+(n-1)a_{n-1}+(n-2)a_{n-2}\\
&+\cdots+3\times a_3+2\times a_2+1\times a_1\}\\
&\quad-na_1-(n-1)a_2-(n-2)a_3\\
&\quad-\cdots-3\times a_{n-2}-2\times a_{n-1}-1\times a_n
\end{aligned}
\tag{49}
$$

ここで，平均差を誘導するときに

$$
\begin{aligned}
T=2\{&na_n+(n-1)a_{n-1}+(n-2)a_{n-2}+\cdots\\
&+3a_3+2a_2+a_1\\
&-na_1-(n-1)a_2-(n-2)a_3-\cdots\\
&-3a_{n-2}-2a_{n-1}-a_n\}
\end{aligned}
\quad (5)\,[再掲]
$$

を得たことを想起する。

(49)式に(5)式を代入すれば，(46)式の右辺は

$$(n-1)A_n-2\sum_{i=1}^{n-1}A_i=\frac{T}{2} \tag{50}$$

となる。

他方で，すでに導出した $n(n-1)$ 個の強度差の総和をあたえる

$$T=2\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i) \quad (7)\,[再掲]$$

の両辺を 2 で割れば，

$$\frac{T}{2}=\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i) \tag{7'}$$

となる。

(7')式と(50)式から

$$
\begin{aligned}
&\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i)\\
&=(n-1)A_n-2\sum_{i=1}^{n-1}A_i
\end{aligned}
\quad (46)\,[再掲]
$$

となる。よって，

(44)式の分子＝(45')式の分子

が証明された。

題意により，(46)式が成立することが証明されれば，

$$R=\frac{\dfrac{2}{n(n-1)}\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i)}{\dfrac{2A_n}{n}}$$

$$=\frac{\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i)}{(n-1)A_n} \qquad (44)［再掲］$$

によって，集中比 R を定義することができる。このことはとりもなおさず，

$$R=\frac{\Delta}{2M_n} \qquad (43)［再掲］$$

によって R が定義されることを意味する。

（証明終わり）

＊　＊　＊　＊　＊

以上要するに，全部で $n(n-1)$ 個ある強度差 (a_j-a_i) ［ただし，$a_j \geqq a_i$，$j>i$］ にかんする平均差を

$$\Delta=\frac{2}{n(n-1)}\sum_{i=1}^{\frac{n+1}{2}}(n+1-2i)(a_{n-i+1}-a_i) \qquad (4)［再掲］$$

とし，強度の相加平均を M_n とするとき，集中比 R は

$$R=\frac{\Delta}{2M_n} \qquad (43)［再掲］$$

によって再定義され，R は平均差と強度の相加平均の 2 倍との比率であるとも規定されることなる。また，(43)式より

$$\Delta=2M_n\cdot R \qquad (51)$$

となり，M_n と R が分かれば，平均差 Δ をもとめることができる。

む す び

本稿における考察は以下のように要約できる。

(1) ジーニ係数は，平均差 Δ を用いても定義できる。1936 年のコールズ委員会研究集録では，分布における集中の測度としてのジーニ係数(集中比)は $R=\frac{\Delta}{2M_n}$ であたえられるという趣旨の叙述が見られる（ただし，M_n は強度の相加平均)。このことから，平均差によるジーニ係数の定義は 1936 年になされたと考えられることもある。しかし，1914 年に初めてジーニ係数が「集中比」として定式化されたとき，すでに平均差による定義が行われていた。

(2) 理論史のうえでは，平均差の概念は，ガウスやハーゲンの名と結びついて発展した 19 世紀中葉の（とくに観測精度の確定を課題としたヨルダン，フォン・アンドレ，ヘルメルトの）誤差論（観測値結合論）研究に淵源する。

(3) 平均差概念にかんする先行研究の存在をジーニが知っていたことは明らかである。ジーニは先行研究を咀嚼してはじめて平均差の研究が完結するという趣旨の主張を述べてはいるが（p.58)，それは彼自身の課題でもあった。この課題の解明については，なお検討を要するが，ジーニは強度間格差を計測する必要から平均差の概念に到達した。

(4) 平均差は強度差の相加平均である。そのため，平均差をもとめるには強度差の総和をもとめておく必要がある。したがって，それには強度差の概念が形成されていることが前提となる。この強度差はヨルダンの観測差の概念にほぼ対応している。ヨルダンは，その個数が $\frac{n(n-1)}{2}$ であると規定

した。その後，フォン・アンドレによって観測差の総和をあたえる計算式が提示された。他方で，ジーニも，フォン・アンドレと基本的には同一の計算式を誘導している。このことから見ると，たとえジーニ理論が先行研究から独立して展開されたとしても，平均差をもとめる途中で誘導される強度差の総和かんする計算式をそのものとして見れば，そこにはジーニに独創性はない。

(5)　強度差の総和を強度差の個数で除して，その相加平均をもとめるとき，ジーニと19世紀誤差論とでは，合計をもとめるべき強度差の個数に違いがあり，誤差論ではその個数がジーニの個数の1/2になっている。

(6)　平均差という用語はすでにヨルダンが1872年に用いている。しかし，この平均差は誤差の平方平均の平方根の$\sqrt{2}$倍であり，ジーニの平均差（強度差の相加平均）とは異なっている。

(7)　ヘルメルトはジーニの平均差$\varDelta$と

$$\vartheta=\frac{1}{\sqrt{2}}\varDelta$$

という関係にある「平均誤差」ϑを誘導して，精度の確定というヨルダン以来の課題に解答をあたえようと試みた。

(8)　ジーニが対象とした社会科学の研究では，相加平均を真値（もしくはその近似値）とみなす誤差論の研究成果がそのまま有効に機能するとは限らない。このために，ジーニは，個体のもつ数量的規定性（ジーニのいわゆる強度）がそのものとして意味をもつ（たとえば，所得のような）社会経済現象を分析するための変動性指数のひとつとして，平均差にその意義を見出した。

(9)　ジーニは二種類の平均差を定義した。一方は「平均差」$\varDelta$であり，他方は「重複平均差」$\varDelta_R$である。今日の表記法ではそれぞれの平均差（mean difference）は

$$\varDelta=\frac{1}{n(n-1)}\sum_{i<j}2\times|a_i-a_j|$$
$$(i=1,2,\cdots\cdots,n-1;\ j=i+1,i+2,\cdots\cdots,n)$$

および

$$\varDelta_R=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}|a_i-a_j|$$

と書かれる。十分大きなnについては$\varDelta\fallingdotseq\varDelta_R$となって，両者に違いはない。しかし，個体の強度差（強度間格差）の平均を計測するという，平均差の本旨から言えば，あえて同一世帯の所得を組み合わせて，その差（格差）を計算することは実質的な意味をもたない。ジーニが用いたように，同一の強度どうしの強度差a_i-a_iを取り除いて計算される平均差$\varDelta$の方が，どちらかと言えば望ましいと考えられる。ツーバーは，重複平均差$\varDelta_R$に較べて$\varDelta$の方が「理にかなっている」と主張しているが，彼の見解の根拠をそのように解釈したい。

(10)　ジーニは1912年論文で平均差$\varDelta$と重複平均差$\varDelta_R$と2つを区別した。しかし，事柄の性質を明確にするために，$\varDelta$には「非重複平均差」という呼称がふさわしい。本稿では，一度でも強度差を計算した強度の組はそれ以降の計算過程では除外し，その上でもとめた平均差を「完全非重複平均差」$\varDelta_P$と名づけ，「非重複平均差」$\varDelta$と区別した。この場合の強度差の個数は$n(n-1)/2$となり，非重複平均差（ジーニの平均差）$\varDelta$の場合の個数$n(n-1)$の1/2になっている。ただし，$\varDelta_P$をもとめるときの強度差の総和は$\varDelta$をもとめるときの強度差の総和の1/2となっているので，$\varDelta=\varDelta_P$となる。また，十分に大きなnについては$\varDelta=\varDelta_P\fallingdotseq\varDelta_R$となる。誤差論における研究を踏まえ，かつ平均差の主旨に照らして見ると，$\varDelta_P$をもって平均差の定義式と見なすが望ましいと考えられる。

⑾　ジーニは平均差（非重複平均差）$\varDelta$ を用いて集中比を再定義した。それにもとづけば，

$$\varDelta = 2M_n \cdot R$$

によって集中比 R から逆に個々の強度間の平均的な格差（乖離）（平均差 $\varDelta$）を計測することができる。これによって，集中比（ジーニ係数）はローレンツ曲線という視覚的手段に加えて，所得分布をさらに明証的に計測する指標となる。

以上に要約されるジーニの研究は，①集中比を含む変動性指数の体系化[49]，②所得分布研究への集中比の応用[50]を誘発した。

本稿では集中比との関連に限定して平均差概念をとりあげたために，体系をなすジーニの平均差理論の全体像を明らかにすることはできなかった。そればかりか，19 世紀後半の誤差論研究者（とりわけ強度差の個数をもとめたヨルダン，強度差の総和をもとめたフォン・アンドレ，そして，ヨルダンとフォン・アンドレの研究成果を踏まえて強度差の相加平均をもとめたヘルメルト）の所説とジーニの平均差理論との関連の考察についても示唆の域を出るものではない。残された今後の課題は重く大きい。

49）Pietra (1915).
50）Ricci (1916).