所得分布に於ける不平等度と、その推計

―附・北海道新得・芽室・美瑛・中富良野に於ける農家所得の実態調査―

京野禎一

目次

§1 問題の所在

所得に関する研究はおよそ次の二つの面に分けて考えることが出来る。その一つは主として所得の大きさに関聯した問題であり、そこでは所得の意義が問はれ、その大いさの確定等が問題になる。その第二は所得の分布或は配分の問題であり、こゝでは分布函数の確定の問題、分布函数の時間的・歴史的な安定性、或は場所的・社会的安定性の問題があり、更にそれらと密接な関聯を持つ所の分配に於ける平等不平等に関する問題がある。而してこの所得の分析が単に所得のみの問題としてではなく、例へば消費とか貯蓄従つて投資とか租税という様な他の諸種の経済的分析との関聯の下に取上げられる時には、それは後者の立場即ち如何に分配されているかとの立場から捉えられる方がより大なる意義を有するものと考えられる。

私はこゝで所得の配分に関する一つの考察を試みたいと考へるのであるが、それは所得分布に於ける不平等度に関する考察である。勿論この点に関しては従来より多くの研究を見るのであるが私の当面の問題の一つは、それらの所得分布に関する分析を検当して、最も妥当であると考える不平等度測定の方法に関説することであり、更にそれらの研究特にその計量的な研究に対して、次の二点を指摘することによつて、私のもう一つの問題点を誘導したいと考えるのである。

1. 従来所得分布の測定に於ては、主として諸種の所得税の統計資料を使つていた様であるが、この点所得税賦課に於ける所得額決定の信頼性を考えるならば、我々が分析せんとしている対象集団の真の所得の配分状態を知らんとする限り、資料は実際の所得額の調査により獲得するという様な方向に進まねばならぬと考える。而して吾々はかゝる資料を対象の集団全数について獲得することは現実に於ける諸種の制約から殆んど不可能であると考えられる。従つてその際五%とか一%とかのサンプルを抽出して、それについて実際の所得額を精査し、以て標本に於ける分布状態を把え、それにより全体の分布状態を推計するという如き方法を取るのが最も合理的であると考える。

2. 吾々が所得の配分を問題にする時は、或一時点に於ける配分状態を知ることよりも、むしろ時の経過と共にその分配状態が如何に変遷して行くかを捉へることの方が必要なのであつて、そのためには出来る限り短い間に、連続的に分布状態を把握し、且その結果を迅速に知ることが望ましいのである。そのためにも、前述の如き標本の分布状態に基いて集団全体の所得の分布状態を推測する如き方法が最も有効であると考える。

以上の二点の反省によつて、近代統計学の提供する統計的推定の方法を援用して、所得分布の不平等度係数の推計の問題を考察せんとするのが私の第二の問題である。

§2 所得分布に於ける不平等度

所得分布の不平等度を取上ぐるに際して、先づ考えねばならぬ事は、吾々は、その係数的把握のために如何なる測定法を採用するかの問題である。そのために従来から用いられて来た不平等度係数の主なるものを概観し、その妥当性を検当せねばならぬだろう。所でそれらの係数としては、周知の如く、パレートの係数　ヂニの係数　ヂブラの係数　ローレンツの係数等がある。以下之等の諸係数を検当せんとするのであるが、それに先立つて予め次の事柄を考えて置きたい。一般に不平等度とは、各人に分配される所得の大きさの相違を意味する如く考えられる様であるが、勿論それは不平等度と無関係のものではあり得ない。而し乍らそれのみで不平等度を何時でも正確に把え得るものではない。不平等度とは、分配されるべき総額の如何なる部分を各々の分配分が占めているか、いはゞその配分比の問題なのである。従つてそこでは所得の単純度数分布に於ける非対称度即ち歪度の影響をも問題にせねばならぬのである。即ち単純度数分布に於て、高額所得者が多数を占め、分布が高額所得層の方に歪んでいるか、或は又分布が低額所得層の方に歪んでいるかによつても、その不平等度の

大きさは異つて来るからである。勿論各人の所得の大きさがすべて等しい時には分配の立場からも完全に平等であると言えるだろう。而して各人の所得に大小の相違があればある程、一般的には不平等度が大になると考えて良いのであるが、所得の大小の相違が不平等度に比例するという如き一義的な解釈は出来ぬのであつて、一方が一定である時のみ、他方の大小で不平等度の大小は決つて行くのである。この事は一々例証する迄もなく明瞭なことである。而して現実にこの両者を同時に考慮し得る如き方法が不平等度を正しく捉える方法であると考え得るのである。予めかゝる観点を明瞭にして置いて前述の不平等度の諸係数の検当を簡単に考察して見たい。先づパレートの係数から始めることにする。

周知の如くパレートの不平等度係数とは次の如きものである。今χを所得の大きさnをx以上の所得を有する人員の累計、α, Aを夫々統計的に定まる常数とすると之等の量の間に次の関係があるとパレートは考えたのである。

$$n=A\cdot\chi^{-\alpha}$$

従つて両辺の対数をとると

$$\log n=\log A-\alpha\log\chi$$

この直線の勾配を示すαを以て不平等度を表すものと考えたのである。而し乍らパレートのこの分布は高額所得層並に低額所得層に対しては当はまらず、実際にこの法則を利用する場合には、この両端の層を除外して考えねばならぬ点と、パレートの場合歪度をも同時に考慮しているが、所得の大きさの相違が大なる時には歪度も又大なる事を予想しているのであつて、両者が自由に変化する場合に関しては妥当性を欠くものと考えられる点とを指摘せねばならぬ。

次にヂニの不平等度係数であるが、それは次の如きものである。χを所得の大きさ、nをχ以上の所得を有する人員の累計、sをχ以上の所得金額の累計、δ, cを統計的に決定される常数とすると、次の関係があるとするのである。

$$n=\frac{1}{c}s^{\delta}$$

従つて　$\log n=\delta\log s-\log c$

この直線の勾配δを以て不平等度を示す係数と考えたのである。この法則はパレートの場合と異つて、所得分布の全範囲にわたつて無制限の適用性を有し、且δはαより敏感に変化する性質を持つている。即ちδとαとの関係は

$$\delta=\frac{\alpha}{\alpha-1}$$

であるが、$\alpha=2$ の時 $\delta=2$ であり、$\alpha=1$ の時 $\delta=\infty$であつて、αが2から1迄変化する間に、δは2から∞迄変化する分である。而し乍らδのこの急激な増加に於て、その増加の解釈に明瞭さを欠く懸念がある。

次にヂブラの場合であるが、元来所得の単純度数分布に於ては低額層の方に強度の歪を見せることが知られているが、ヂブラの場合、横軸を対数に目盛れば大体正規分布をすると考えるのが、その根本的立場である。従つてヂブラの考えた法則は次の如きものであつた。

今所得額をx、所得人員をyとすると、

$$y=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\log^x-g}{\sigma}\right)^2}$$

が成立するとするのである。（但し、gは幾何平均である。）こゝで

$$-Z^2=-\frac{1}{2}\left(\frac{\log^x-g}{\sigma}\right)^2$$

とおくと

$$Z=\frac{\log^x-g}{\sqrt{2}\,\sigma}=\frac{1}{\sqrt{2}\,\sigma}\log^x+\frac{-g}{\sqrt{2}\,\sigma}$$

更に $\frac{1}{\sqrt{2}\,\sigma}=a$, $\frac{-g}{\sqrt{2}\,\sigma}=b$ とおくと

$$Z=a\log x+b$$

而して、今この半対数正規曲線の面積をRとすると

$$R=\int_Z^{+\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{-Z^2}$$

このRの総計は1であるが、之を千分の1の間隔に区切つて、それに対応するzの値をヂブラは計算しているのである。そこである所得分布を考えた時には、その階級別所得人員の累計の構成比がRに相当するであろう。従つてヂブラの数表からそれに対応するZを求めて、図表を描けばほゞ直線になると言うのである。そうしてこの勾配aを以て不平等度係数と考えたのである。而し乍ら実際には、Zはかなり彎曲して居り、従つてもとの度数分布は半対数正規曲線をなさぬ様である。それ故に直線を当はめるにはかなり無理がある様であり、強いて直線の当はめをなさんとせば次の如き修正を要するであろう。即ち

$$Z=a\log(x-x_0)+b$$

而し乍らこの時は原資料の不平等度の測定でなくして、修正された資料の不平等度の測定となつて了うのである。更に不平等度の係数は $a=\frac{1}{\sqrt{2}\,\sigma}$ であつて、むしろ直接σを用いた方が理解し易い様であり、又この場合は当然歪度を零と考えているのであつて、所得の大小に於ける相違のみが不平等度決定の要因となつている様である。

最後にローレンツの係数を考えて見よう。xを所得x以下の所得人員の累計、yを所得x以下の所得金額の累計とし、且その各々の百分比を夫々X. Yとする。対応するX. Yを座標に持つ点を平面上にとつて行き、それらを結んで出来る曲線がローレンツの曲線である。そこでローレンツの不平等度係数としては、平等配分を示すY=Xなる直線とローレンツの曲線とによつて囲まれた部分の面積をλとすると、このλの大きさを以て表はさんとするのである。而してこの面積の計算に於ては、その曲線の函数形が決定すれば勿論、函数形が決定出来ぬ場合でも、シンプソンの近似計算を行うか、或はプラリメーターで機械的に測定出来るのである。然し乍らこのローレンツの曲線に関しても、λのみでは所得の大小の相違を表し得るが、歪度の大小に関しては充分にその相違を表はし得ないと考えられている。勿論歪度大となれば、不平等度も一般には大になると言う程度の表現はλにも含まれていると考えるのであるが、λの大きさがほゞ同じであつても分布の

形状を異にする場合があり得るのである。この点の解決が実は私の一つの問題であつて、それは曲線或は函数式の更に細かな検当によつて、歪度の差違をも考慮に入れた不平等度の分析が出来ると考えるのである。

以上の不正等度係数の概観を終へて、私は所得分布に於ける不平等度係数としては、ローレンツの曲線（単にλのみでなく）によるのが差当り妥当であるとの結論を得るのである。

更には不平等度測定に於ける労力とか、その意味の理解の難易とか言つた点をも考慮に入れるならば、一層この感を強めるものである。

§3 ローレンツの曲線による不平等度

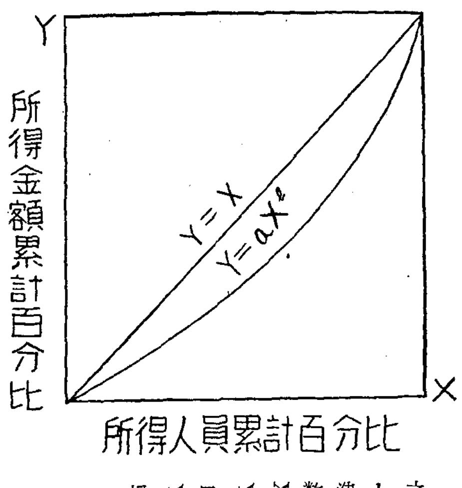
ローレンツの曲線による不平等度を考える時、単にその面積λのみが問題であるならば、前述の如く曲線の函数形は確定せずとも良いのであるが、更に歪度の考察をも加味した分析を行うためには、何らかの方法で曲線の函数形を決定し、λが同一の場合でもその分布曲線の状態でその所得分布の差異を見て行かねばならぬのである。而してローレンツの曲線は下図にも見られる如く、経験的ではあるが近似的に一定の傾向をもつた曲線であることが知られて居り、その函数形は

$$Y=X^{b}$$

としてヂニの法則より導出されている様である。然し乍ら私はもつと一般化し、その函数形を次の如き形であるとするのが妥当であると考える。即ち

$$Y=aX^{b}$$

之である。aｂは統計的に決定される常数であり、X＝0の時はY＝0 X＝100の時は、Y＝100の前提の下で、0＜a＜1、1＜bの範囲で種々なる値を取ることが



出来、従つて曲線の形はその彎曲度、傾斜の様子等であらゆる場合を尽し得ると考へる。且a＝1 b＝1の時が平等配分の場合であつて函数形はY＝Xとなるであろう。かくしてローレンツの言う不均等度λは次の如くして計算出来るのである。

$$\lambda=\int_{0}^{100}(X-aX^{b})dX$$

曲線が平等分配線Y＝Xより離れゝば離れる程分配が不平等になるのであるから、λが大になればなる程不平等なることを示す。従つて実際の資料により、この常数a、bを決定して、上述積分計算により数値的にλを得ることが出来る。而し乍らその前

に注意されねばならぬ点は

第一に元来所得分布の不平等度に関しては、ある所得配分の理想的状態を考えて、現実の分布がその理想的配分状態から如何に乖離するかが問題なのであつて、ローレンツの場合の理想的配分状態としては平等配分が考えられているのであり、所得総額の一〇%、二〇%、三〇%等々が夫々一〇%、二〇%、三〇%、等々の人々によつて所有される状態を考えているのである。この点に関し理想的配分状態を如何に考えるかゞ問題になるであろう。従来迄は平等配分を以て理想的配分状態と考えられて来たのであるが、最近では如何なる配分状態がその経済集団全体の厚生を大にするかと言う様な觀点から考えられている様である。従つてその場合には分配は必ずしも平等配分の視点から考えられては居ない。然し乍らこの様な状態を実際に確定することは仲々困難であり、ローレンツが考えた平等配分を一応こゝで理想的配分状態と考えても、大して妥当性を欠くものではないと考える。且その理想的な配分状態が如何なる意味にてでもあれ確定した時には、その不平等度の計測としては、この場合と原理的には同じであると考え得るのである。

第二には、先にも述べたる如くλのみでは所得分布の不平等度を完全に表現するものではなく、単純度数分布に於ける歪度をも考慮に入れた分析が行はれねばならぬと言う点である。

そこで以下に於てこの点の考察を進めて見たいと考える。そのために今、実際にローレンツの曲線の函数形が、$Y=aX^b$ と確定したものとしておく。今この曲線に関してその不平等が所得人員

の何%位の所でその最大を示すかを見よう。そのためには

$$\phi=X-aX^b$$

なる函数を考えこの新な函数ϕの最大値を求めれば良いことになる。そこで第一次並に第二次の導函数を導くと

$$\frac{d\phi}{dX}=1-a\cdot b\cdot X^{b-1}$$

$$\frac{d^2\phi}{dX^2}=-a\cdot b\cdot(b-1)X^{b-2}$$

現在の所 $b>1$ と考えてよいのであるから

$$\frac{d^2\phi}{dX^2}<0$$

故に $\frac{d\phi}{dX}=0$ と置いて求まる所のXで極大（即ち最大）が起きるのである。次にその様なXを求めて見よう。

$$\frac{d\phi}{dX}=1-a\cdot b\cdot X^{b-1}=0$$

$$\therefore\ X^{b-1}=\frac{1}{a\cdot b}$$

$$(b-1)\log X=-\log a\cdot b$$

$$\therefore\ \log X=\frac{-\log a\cdot b}{b-1}$$

之からXの値が求まる。今かゝる不平等を最大にするXの値をX_mとすると、このX_mに於ける勾配は1であつて X_m 以下の所得人員に於ては、人員の単位増加に対する所得金額の増加割合は1以下であり、X_m以上の所得人員に於ては、人員の単位増加に対する所得金額の増加割合は1以上である。従つてこのX_mの値

の大小によつて、一応その分布に於ける歪度に相当する分析が可能である。即ちX_mが大なる程歪度は一般的に大となるであろうからである。而し乍らこの点を更に細かに見るならば、X_mがほゞ同じ値になつたとしても $X=X_m$ なる時のYの値をY_mとすると、Y_mの値の大小によつて又その配分の平等、不平等の様相は異つて来るのである。即ち同じX_mに対してもY_mの値が小なる程、単純度数分布に於ける歪度は大となるものと考え得る。この点をも更に考慮に入れるために次の如き考察を加える。即ち単位百分比の所得人員の増加に対して、所得金額は百分比に於てどの位の割で増加するかをX_m以下の所得階層とX_m以上の所得階層とに於て計算することである。その平均の数を夫々 μ_1, μ_2 とすると、それは次の如くして求め得るであらう。勿論前述せる如く $\mu_1 < 1$ $\mu_2 > 1$ である。

$$\mu_1 = \int_0^{X_m} Y'\,dX \Big/ X_m$$

$$\mu_2 = \int_{X_m}^{100} Y'\,dX \Big/ (100-X_m)$$

但し　$Y' = \dfrac{dY}{dX} = a.\,b\,X^{b-1}$

而して μ_1 が小なれば小なる程、μ_2 が大なれば大なる程歪度は大になると考えられる。そこでこの μ_1 μ_2 の両者を併せ考えるため

$$\mu = \mu_2 : \mu_1 = \frac{\mu_2}{\mu_1}$$

なる値を考える。この μ の値が大になれば大になる程、単純度数分布に於ては歪度が大になり、従つて一般的には不均等度が大になると考え得る。若しも或集団に於ける所得の単純度数分布が正規分布であつたとすれば、その時は μ_1 と μ_2 は逆数関係にあり μ は1となる筈である。$\mu<1$ の時は高額層に歪み、$\mu>1$ の時は低額層の方に歪んで居り、その大きさ大なれば大なる程、強度の歪みを見せることになる。従つて λ 略同一の時は μ の差を問題にすれば、その所得分布の不平等度の判定可能であり、λ が異つていれば μ を見る迄もなくその不平等度の差違を一般的に結論し得る。

　ともかくもローレンツの曲線による不平等度の測定としては、この λ 並に μ の両面から考察されることが望ましく、λ も μ も大きくなればなる程、不平等度を大にするのであるから、λ と μ との積によつて、その分布の不平等度をほゞ正確に判定し得ると考えるのである。

　以上でローレンツの曲線に関する不平等度の分析を終へたのであるが、尚更に之に関して附加えるならば、先づ第一に理想的配分の函数形　$Y=X$　では　$\dfrac{dY}{dX}=1$　$\dfrac{d^2Y}{dX^2}=0$である。従つてローレンツの曲線の函数形の第一次導函数 $\dfrac{dY}{dX}=a.b.X^{b-1}$ のグラフの勾配が小になればなる程、換言せば $\dfrac{d^2Y}{dX^2}=a\cdot b\cdot(b-1)X^{b-2}$ の積分値が小なれば小なる程、理想配分状態に近くなるので、その値即ち

$$\int_0^{100} a.b.(b-1)X^{b-2}dX$$

は　λの代りに用ひ得ると考える。

第二には、ローレンツの曲線に於て平均に於ける、所得人員の単位百分比の増加に対する所得金額の百分比に於ける増加の割合即ち

$$\frac{\int_0^{100} Y' \, dX}{100} = \frac{a.b \int_0^{100} X^{b-1}dX}{100}$$

の大小によつて　μ の大きさの大凡の傾向を見得ると考える。

§4　ローレンツ曲線に於ける不平等度の推計

次に私は所得分布に於ける不平等度の推計の問題を、ローレンツの曲線に関して取扱う積りである。そのためには当然ローレンツの函数形　$Y=aX^b$　に於ける a, b の数値の決定方法を先づ考えねばならぬ。それは周知の如く曲線当はめの問題である。而し乍らこゝで推計の問題をも同時に考慮した時に、それを容易ならしむるため、ローレンツの函数の両辺の対数をとり一次式に変形して、a, b の決定並にその推計を行うことにする。従つてその一次式は次の如くになる。

$$\log Y=\log a+b \log X$$

この時言う迄もなく logX, logY を変数と考えると、そのグラフは直線である。今わかり易くするため　$\log X=x$.　$\log Y=y$　と置換して、一次式は　　$y=A+BX$

但し　$A=\log a$,　　$B=b$

X, Y, x. y の変域は

X, Y, が 0→100 の時夫々 x, y は 0→2

である。従つてA、Bが決定されゝばよいのであつて、そのためには x の上の y の回帰直線の方程式を求めることであり、最小自乗法を用いて

$$\varphi = \sum_{i=1}^{n} (y_i-y)^2 = \sum_{i=1}^{n} (y_i-A-Bx_i)^2$$

を最小ならしめる如くA、Bを定めればよいのである。そのためには

$$\frac{\partial\varphi}{\partial A} = -2\sum_{i=1}^{n} (y_i-A-Bx_i)=0$$

$$\frac{\partial\varphi}{\partial B} = -2\sum_{i=1}^{n} x_i(y_i-A-Bx_i)=0$$

なる如くA、Bを定めるのである。上式から

$$\sum_{i=1}^{n} y_i - nA - B\sum_{i=1}^{n} x_i = 0$$

$$\sum_{i=1}^{n} x_i y_i - A\sum_{i=1}^{n} x_i - B\sum_{i=1}^{n} x_i^2 = 0$$

$$\therefore A = \frac{\Sigma x_i^2 \, \Sigma y_i - \Sigma x_i \, \Sigma x_i y_i}{n\Sigma x_i^2 - (\Sigma x_i)^2}$$

$$B=\frac{n\Sigma x_i y_i-\Sigma x_i\Sigma y_i}{n\Sigma x_i^2-(\Sigma x_i)^2}$$

而してかゝるA、Bが決定された後は $B=b$, $A=\log a$ の換算によつてa、bが得られ、従つて $Y=aX^b$ なる函数は確定する。

次にかゝる函数が標本のそれであるとして、対象集団全体についてのローレンツの函数の推計方法について述べるのであるが、その際標本として得られる X_i Y_i の値従つて x_i y_i の値は言うまでもなく任意標本の値でなければならぬ。以後変数は x, y 迄変換したものとする。

今対象集団全体について x, y の系列を考え、そこからn個の標本値 x_i, y_i $(i=1, 2, \cdots\cdots, n)$ を得たものとし、かゝる変数 x_i, y_i に関して前述の回帰直線の係数A、Bが決定されたとする。その値を今便宜上 A_s, B_s とする。而る時に、推計の順序として先づ集団全体に関する回帰係数Bを推計し、而る後にAを推計することにする。即ちBの推計では回帰係数の推定方法となり、Aの推計に関しては、こゝでは簡単にするためにBが決つた時、その回帰直線は x_i, y_i の平均 $(\bar{x}, \bar{y})$ を通るのであるから、その条件を入れることによつてAを決定する方法を用いることにする。

そこで先づBを推定する方法を考えたい。

今 x, y の系列の中から大きさnの無作為標本を繰返し抽出する場合を考えるのであるが、その際 x_i に関しては最初の一組で得た標本の値を何時でも取らせるものとする。この x_i に対する制限は最初に大きさnの任意標本が取られた後には x_i は変数でないことを意味する。事実最初の標本を取る時でも x_i は任意でなくとも良いのである。そうしてかゝる抽出の繰返しに於ては y_i は $y=A+BX$ で示される集団全体に於ける回帰直線の周りに等しい分散 σ^2 を以て正規に分布すると考えてよい。即ち x_i は各抽出に於て固定されているが、y_i は各抽出毎に変つた値をとり、上述の如き分布をする変数と考えて良いのである。而してかゝる抽出を繰返し、その抽出毎に B_s を求めるとその B_s は如何なる分布をするか考えて見る。前に示した B_s を求める式

$$B_s=\frac{n\Sigma x_i y_i-\Sigma x_i\Sigma y_i}{n\Sigma x_i^2-(\Sigma x_i)^2}$$

は、次の如き形に書き変えることが出来るであろう。

$$B_s=\frac{\Sigma(x_i-\bar{x})y_i}{n\Sigma(x_i-\bar{x})^2}$$

即ち後の式を展開することによつて、容易に前の式を得るであろう。そこで今の場合は x_i, $\bar{x}$ は常数と見做して良いのであるから結局 B_s は y_i の一次式であり、y_i は互に独立に正規分布する変数と考よいるのであるから B_s も又正規分布をする筈である。而してその分散も又、以上の前提の下で容易に次の如く求め得る。

$$\sigma^2_{Bs}=\frac{\sigma^2}{\Sigma(x_i-\bar{x})^2}$$

従つて今

$$u=\frac{B_s-B}{\sigma_{Bs}}=\frac{(B_s-B)}{\sigma}\sqrt{\Sigma(x_i-\bar{x})^2}$$

と置くとこの u は平均値零、分散 1 の正規分布に従うのである。

更に又次の如き値を考えて見る。

$$v^2 = \frac{\Sigma(y_i - y_i')^2}{\sigma^2}$$

（但し、y_i' は標本回帰直線による推定値）

而る時はこの v^2 も又自由度（n−2）なる χ^2（カイ）− 分布に従うことが知れている。そこで

$$t = \frac{u\sqrt{n-2}}{v}$$

なる t を考えると、それは自由度（n−2）なるステューデントの t− 分布に従うのである。以上の事柄は詳細な経過を省略したが数理統計学の良く証明する所である。この t− 分布の性質を使つて B の推定を行はんとするのであるが、この t の値をもとの変数にもどして見ると

$$t = \frac{\frac{(B_s - B)}{\sigma}\sqrt{\Sigma(x_i - \bar{x})^2}\,\sqrt{n-2}}{\frac{\sqrt{\Sigma(y_i - y_i')^2}}{\sigma}}$$

$$= (B_s - B)\sqrt{\frac{(n-2)\Sigma(x_i - \bar{x})^2}{\Sigma(y_i - y_i')^2}}$$

となる。

t− 分布とは、最初ステューデント（ゴセットの筆名）が考え、後フイッシャーが幾何学的な考えから、その分布の函数型を考え

たもので、自由度小る間は正規分布よりも少し大きいひろがりをもつた分布で、自由度が大きくなるに従つて正規分布に近づき自由度30以上なる時は殆んど正規分布と見なして良い様である。更にフイッシャーは、この t− 分布に関して $|t| > t_\varepsilon$ なる確率が ε なる様な表を計算して居り、それは各々の自由度に対して信頼度が ε なる如き t の値であり、t がその値を越える確率が $1-\varepsilon$ である様なものである。従つて或自由度（之は標本数が決れば決定する）に対して信頼度 ε なる如き t_ε をフイッシャーの表から見出し

$$|t| > t_\varepsilon$$

とおけば B の含まれる範囲が信頼度 ε の下でかくかくであると立言出来るのである。即ち

前記の $\sqrt{\frac{(n-2)\Sigma(x_i - \bar{x})^2}{\Sigma(y_i - y_i')^2}} = \nu$ とおくと

$$|t| = |(B_s - B)\nu| = |B_s - B|\nu > t_\varepsilon$$

$$\therefore |B_s - B| > \frac{t_\varepsilon}{\nu}$$

故に　$B_s - \frac{t_\varepsilon}{\nu} < B < B_s + \frac{t_\varepsilon}{\nu}$

そこで集団全体の回帰係数 B は $B_s - \frac{t_\varepsilon}{\nu}$ と $B_s + \frac{t_\varepsilon}{\nu}$ との間にあり、その時の信頼度は ε であると言える分である。この場合の ε は各実際の問題の性質に応じて、各分析者が決めれば良いのであつて通常は 0.95 とか 0.99 とかが用いられている。従つてそれに対応して、$b_s = B_s$ であるから集団全体の b の値も $b - \frac{t_\varepsilon}{\nu}$ と

$b+\frac{t_\varepsilon}{v}$ の間にあると推定出来るのである。

次にAの推定であるが、それは前にも述べたる如く先に求めた $B_s-\frac{t_\varepsilon}{v}$, $B_s+\frac{t_\varepsilon}{v}$ なる勾配を持つ両直線（今簡単のため、$B_s-\frac{t_\varepsilon}{v}=B_1$, $B_s+\frac{t_\varepsilon}{v}=B_2$ とおく）

$$y=A_1+B_1x$$
$$y=A_2+B_2x$$

が共に（$\bar{x}$, $\bar{y}$）を通るという条件により

$$\bar{y}=A_1+B_1\bar{x} \quad \text{故に} \quad A_1=\bar{y}-B_1\bar{x}$$
$$\bar{y}=A_2+B_2\bar{x} \quad\quad A_2=\bar{y}-B_2\bar{x}$$

として求まる。而る後に $A_1=\log a_1$, $A_2=\log a_2$ なる換算によつて a_1, a_2 は求まり、従つて集団全体の曲線は

$$Y=a_1X^{b_1}$$
$$Y=a_2X^{b_2}$$

の間にあると信頼度εの下で言い得るのである。そこで集団全体としてのローレンツの不平等度に関する分析が、この推定の巾の範囲で可能となるのである。

標本による集団全体の不平等度の分析に於て、かゝる巾を附けたまゝで考案されねばならぬ事は、小数の標本から推測する以上止むを得ぬ事柄であつて、たゞその巾の大きさが客観的に捉えられているので、実際には、その点を考慮に入れて、個々の分布の様相に応じた具体的な分析が可能なのであり、この点単に全数についてその所得を知り、その数が膨大であるため調査粗雑で種々の偏倚を内包し、一般的には信頼性うすく、その誤差の大きさも保証されていない様な分析を行うよりは、はるかに有効であり、又すぐれているものと考えられる。又この推定の巾も小なる程すぐれている事は言をまたぬが、それは種々標本抽出を工夫することによつて狭めて行く事も考え得るのである。

§5 所得分布の不平等度推計の例

最後に以上の所得分布に於けるローレンツの不平等度の分析を過日、本道の新得、美瑛、芽室、中富良野で行つた農家所得の標本調査に適用して考察して見たい。

各町村共標本農家の選択は、任意抽出の一方法である系統的任意抽出法により、その抽出率は五％である。即ち農家名簿の最初の二〇戸の中より全くランダムに一戸選び以下二〇戸間隔に農家を抽出した。以下その分析の大要を述べるのであるが、紙数の制限もあり、詳細な計算過程は一切省略して結果のみをかゝげることにする。最初に各町村に於ける農家総戸数並に抽出戸数は左の如くである。

	新得	美瑛					芽室	中富
		水田	畑	混同	開拓	計		
総戸数	571	525	532	607	400	2064	1466	891
抽出戸数	29	27	27	31	21	106	74	44

但し、新得は畑作地帯として水田農家、開拓農家を除いて 五七一戸

所得分布に於ける不平等度と、その推計

美瑛は農業形態別に見んとして、四つに層別　合計　二〇六四戸

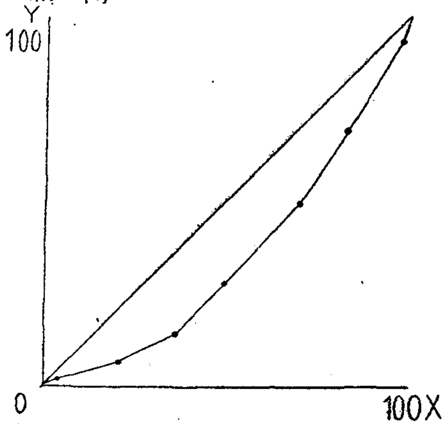
芽室は畑作地帯として水田、開拓農家をおとして　一四六六戸

中富は水田地帯として、畑作、開拓農家を除いて　八九一戸

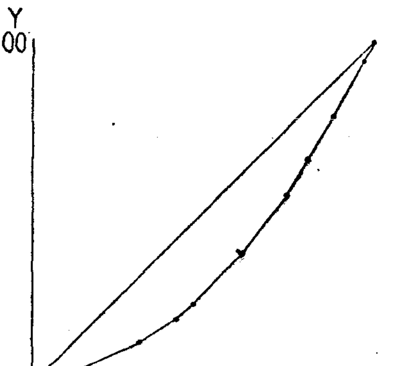
である。

そこで之等の抽出農家について、各町村毎ローレンツの曲線を描いたら下図の如くであつた。

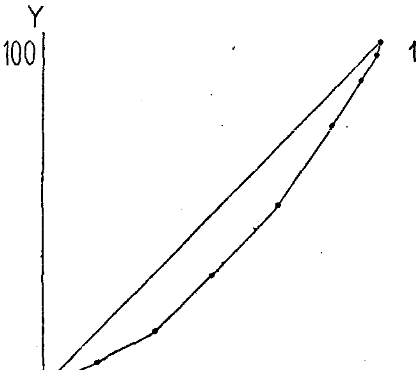
新得



美瑛

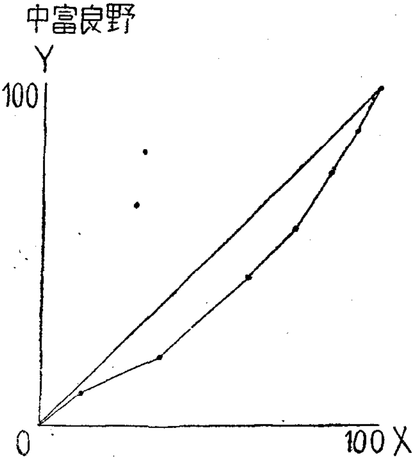


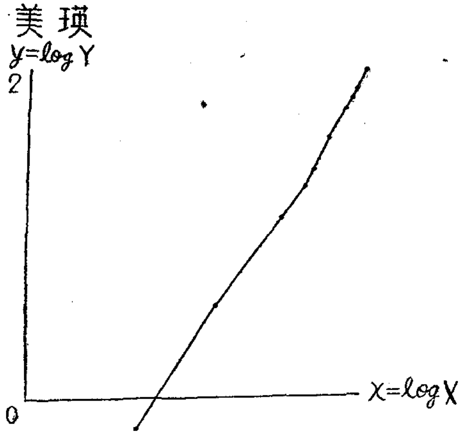
芽室

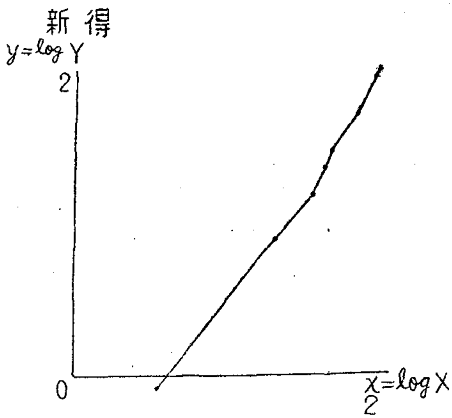


上で見た通り、新得、美瑛、芽室共に殆んど $Y=aX^p$ なる分布をして居り、たゞ中富良野のみ、やゝ低所得層で攪乱を見る。

次にX、Yの対数をとつて、その対数変数でグラフを書いて見ると次の図の如く殆んど直線となり、やはり中富良野のみ目立つた折れ線を見せている。







之等の各直線に当てはめた直線の方程式は次の如くになつた。

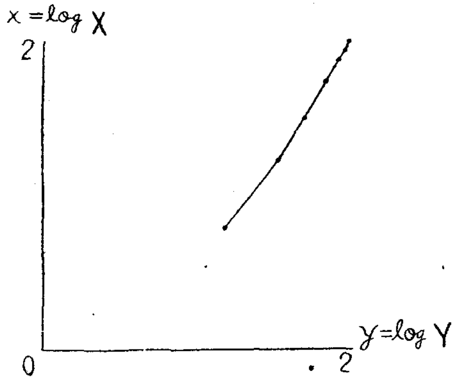
	A_s	B_s	標本による直線の方程式
新得	−0.9763	1.456	$x=-0.9763+1.456y$
美瑛	−1.6207	1.815	$x=-1.6207+1.815y$
芽室	−1.2901	1.653	$x=-1.2901+1.653y$
中富	−1.0100	1.542	$x=-1.0100+1.542y$

従つてローレンツの曲線の函数形は次の如くになる。

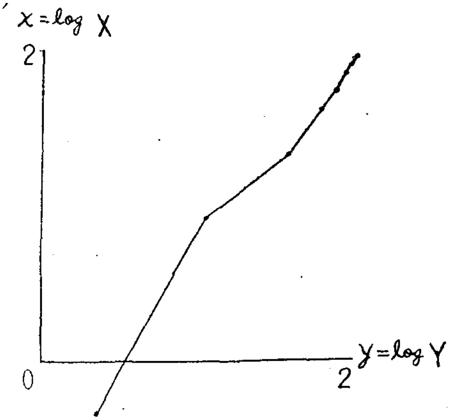
	a_s	b_s	標本による曲線の式
新得	0.10561	1.456	$Y=0.10561X^{1.456}$
美瑛	0.02395	1 815	$Y=0.02395X^{1.815}$
芽室	0.05127	1.6525	$Y=0.05127X^{1.6525}$
中富	0.09772	1.5421	$Y=0.09772X^{1.5421}$

次にこの標本の値によつて、各町村全体の直線の方程式並に曲線の函数形を推計するのであるが、そのために先づ

芽室



中富良野



$\nu = \sqrt{\dfrac{(n-1)\Sigma(t_i-\bar{x})^2}{\Sigma(y_i-\bar{y}_i)^2}}$ を計算し、次にこゝでは信頼度 $\varepsilon=0.99$ として $t_{0.99}$ を t一表からひいて見た。たゞ自由度は新得の場合のみ 29—2 であつて他はすべて30以上なる故、新得では $t_{0.99}=2.771$ で他の場合は共に $t_{0.99}=2.576$ である。従つて $\dfrac{t_{0.99}}{\nu}$ を計算して見ると、次の如くになつた。

	新　得	美　瑛	芽　室	中　富
$\frac{1}{\nu}$	0 0284	0.0183	0.0216	0.0292
$\frac{1}{\nu}\cdot t_{0.99}$	0.07316	0.047	0.0556	0.0752

従つて各町村全体に於ける直線の方程式並に曲線の函数形は、次の如き二つの式の間に含まれることになるのである。

	直線の方程式	曲線の式
新　得	$y=-0.85864+1.38284x$ $y=-1.09382+1.52916x$	$Y=0.13847X^{1.38284}$ $Y=0.080571X^{1.52916}$
美　瑛	$y=-1.54313+1.768x$ $y=-1.69830+1.862x$	$Y=0.028633X^{1.768}$ $Y=0.020031X^{1.862}$
芽　室	$y=-1.19192+1.5960x$ $y=-1.38863+1.7081x$	$Y=0.064281X^{1.5969}$ $Y=0.040867X^{1.7081}$
中　富	$y=-0.89337+1.4669x$ $y=-1.12912+1.6173x$	$Y=0.12783X^{1.4669}$ $Y=0.074282X^{1.6173}$

次に標本に於けるローレンツの不平等度係数λ、並に各町村全体としてのλのある範囲を示すと

	λ		λの百分比（λ/50000）	
	標本値	推定巾	標本値	推定巾
新　得	1488.64	1612.0 1356.54	29.8	32.24 27.13
美　瑛	1370.96	1410.0 1292.9	27.4	28.2 25.86
芽　室	1098.65	1132.6 1065.29	22.0	22.65 21.3
中　富	333.5	550.8 129.0	6.7	11.0 2.6

之によつて見ると、推定の巾を考慮に入れて見ても、新得がその不平等度が最大で、以下美瑛、芽室、中富良野の順となつている。従つて之の値のみで不平等の差異を判定しても良いのであるが尙更に X_m 並にその推定の巾を求めて見ると左の如くであつて、之の値も又λの場合と同様、新得、美瑛、芽室、中富良野の

	標本の X_m	X_m の推定巾
新　得	60.7	75.07〜52.30
美　瑛	46.86	48.65〜45.39
芽　室	43.93	45.32〜42.92
中　富	32.82	36.06〜30.97

順に小さくなつている。

而し乍ら前節にも述べたる如く、より詳細に見るために μ_1 μ_2 μ を計算して見ると次の如くである。

	μ_1		μ_2		μ	
	標本値	推定巾	標本値	推定巾	標本値	推定巾
新得	0.6868	0.7233 0.655	1.4835	1.8331 1.3982	2.16	2.54 2.1
美瑛	0.5506	0.5657 0.5371	1.3964	1.4115 1.3847	2.536	2.5 2.578
芽室	0.605	0.6249 0.5855	1.3095	1.3109 1.3117	2.165	2.08 2.24
中富	0.6636	0.6816 0.6183	1.1644	1.1795 1.1712	1.755	1.73 1.89

之によつて見ると μ の最大なのは美瑛で新得、芽室が之に次ぎ略同じで中富良野が一番小さい。従つて λ 並に μ の両者を同時に考慮するため λ と μ との積を計算して見ると、次の様になる。

	$\lambda\mu$の標本値	$\lambda\mu$の推定巾	
新得	64.303	81.567 56.973	但し λ としては実数を用いず百分比を用いた。
美瑛	69.596	70.5 66.667	
芽室	47.52	47.112 47.712	
中富	11.725	19.03 4.914	

以上を観察して見て、美瑛と新得に於ては大して差違を見ないが、最も不平等度の大なのは美瑛でありそれに次いで新得それから芽室、最も小なのが中富良野という順になつている。この新得と美瑛とは前者が λ に於て大なるに反し後者は μ に於て大となり、従つてこの両者を同時に考えた時には、多少の差はあるが、ほゞ同じ不平等度を示すものと判定して良いと考へる。尚新得に於ては推定に於ける変動の巾が大きいのであるが、之は標本抽出数が30未満なるためで、もう少し標本数を多くして、この変動の巾を縮め得るものである。

法経會論叢（十二集）《正誤表》

執筆者	頁	段	行	誤	正
根手	五		五	形成され）	形成された）
〃	三五		一九	具	且
嘉東	八三		終一〇	北海道財政にもまた……	「もまた」は不要
金田	九〇		二	消減	消滅
〃	九二		終三 一	單一・複数・複合的 基準 との標識	單一・複数・複合的基準 とその標識
京野	一二一	上	四	と 0.99-2.576	と 0.99=2.576
記事	一四三	上	題目	《法経記事》	《法経會記事》
〃	同	中	一五	「宗教硬一つ計画経済に於ける……」	「計量経済に於ける……」
後記	一四四	中	終一	工風	工夫

以上