Lorenz Curveについて

米澤, 治文
東北帝國大学法文學部

出版情報：統計数理研究. 1 (1), pp.53-59, 1941-10-15. Research Association of Statistical Sciences

Lorenz Curve について

會員 米 澤 治 文 (東北帝大法文學部)

(昭和十六年四月卅日受理)

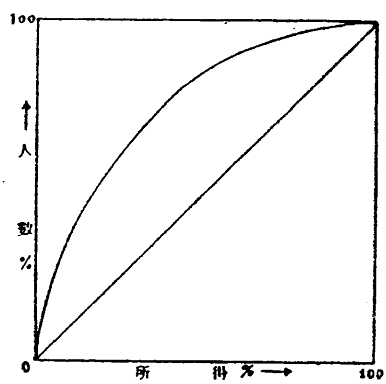
去る四月末，仙臺の統研會員のセミナールに於て報告を行つたところに基いて本稿を草する．Lorenz curve については經濟統計方面以外ではあまり知られてゐない様なので，その紹介の意味で書いたものである．從つて叙述は序論的な部分に限られ，同じ主題についての數理的な展開は別の擔當者によつてなされる筈である．

始めに Lorenz curve について Kurtz and Edgerton の Statistical dictionary の與へてゐる說明を參照すれば，それは次の如くである.[1)]

「財產所得賃銀等の如き變量度數の分布を圖示する一方法．それは本質に於て百分比であらはされた二つの累加的 (two "more than" cumulative) 度數分布である．それでデイタは圖表上から次の如く讀みとられる．『國全體の財產の 60 per cent が全人口の 15 per cent の最も富める者の手中にある』といふ様に．Lorenz curve は平等分布からの可能なる最も極端な乖離の例を自働的に具現せしめるものである.」

今例へば所得の分配に就て觀察して，所得の少額なる者から順次多額の者へ向つて數へて，總人員中の一定割合の者が受ける所得の和が，總人員の受ける總所得の幾何割合であるかゞ知られたとする．此の時直角座標の縱軸にその様な人員割合を，横軸に同じく所得額割合をとつて，その二者の關係を曲線にあらはすならば，これ即ち所謂 Lorenz curve であつて，之によつて所得の各人員間への分布の性向が判定されることとなる．今所得總額の 10% は人員總數の 10% が之を受けるといふ様に，所得人員双方の割合が常に完全に一致するとすれば，上記の様にしてそれを圖示する時に兩軸と恰度 45 度の角をなす直線が描かれる筈であり，かゝる場合は所得分布が正に全く平等な譯である．然るに實際には此の様な完全平等分布は恐らく存在せず，從つて，上記の如くにして描かれた線はかゝる45度の直線と離れたところを通過する．そこでその乖離の程度の如何によつて所與の所得分布の不平等の大小を判定しようとするのである (A 圖參照).

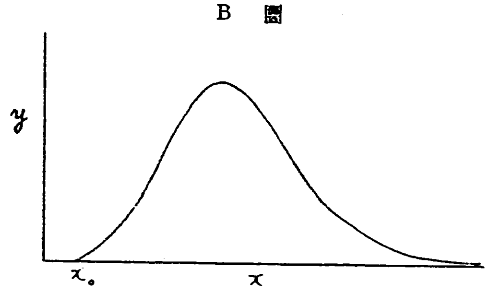
A 圖



かくの如き所得或ひは財產の各人員間への分布の態様如何は一般に統計値の分布の平等不平等を論ずる最も普通の場合である．そしてこの様な場合は一定人員への分布量の集中が問題となるのであるから，これを人的集中現象と呼ぶ.[2)] 此の様な現象に對して Lorenz curve がその適用を見ることは上の例示によつても明かであるが，抑も統計値の分布の不平等乃至集中の如何が論ぜらるべきは必ずしもかゝる人的集中現象には限らない．其他に地理的集中現象の如きものもあり，事物的集中或ひは時間的集中を示す事例も考へられる．けれども從來の例から見ると Lorenz curve の

實際の適用は殆んど凡て人的集中現象に限つて居つて，他種の諸現象には餘り及んで居ない．それは人的集中現象が社會的に最も重要な且つ注目を惹き易い問題であつたといふことが大きな原因に相違ないが，その他にも別の理由が存すると考へられる．それは寧ろ形式的なものであるが，統計材料の取扱ひが簡單であるといふ點である．即ち人的集中現象にあつては統計單位が個人といふ一つ一つ數へ得べき而もその重要性がどれをとつても齊一な個體であるがためである．然るに他の場合には此の點が必ずしも左樣ではないので，それがそれ等の現象と Lorenz curve の間柄を疎遠ならしめる契機となつてゐることを認めざるを得ない．けれどもそれが同曲線の適用不可能を意味するか否かは尙檢討を要すべきであり，後にそれは行はれるであらう．併し兎も角も人的集中現象が Lorenz curve の適用される最も原則的な場合であることは實際上否まれぬところであり，我々が此處で該曲線の本質を少しく究明しようとするに當つても，先づそれは人的集中現象に對して適用されるものとして，矢張り其の樣な設例に基いて進んで行くこととする．

B 圖

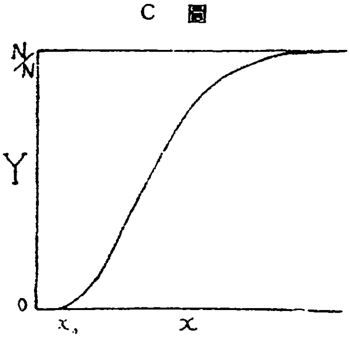


冒頭に引用した文句にもある通り，Lorenz curve は度數分布の一種の累加形態であるといふことは屢々明言されてゐる．然るに一歩進んでそれは如何なる累加形態であるかといふ點になると，必ずしも闡明されては居らない，惟ふにそれは次の如くに解せらるべきであらう．

今此處に x なる所得を有する人員數 y を示した度數分布が與へられたとする（B圖參照）．之はもとより變量（此處では所得）を一定數の階級に分つて，それに對應する度數（此處では人數）を圖示したので，本來度數 polygon を以て表はされるのであるが，充實した統計資料が得られ，階級區分の幅が充分小さくとり得る際には近似的に平滑な曲線を以て表示し得るのである．尙所得の概念上 x は常に正であるとする．

其處で今かゝる度數分布の原形に對して種々なる累加形態を構想して見るに，其の最も普通のものは所謂累積度數分布である．之はB圖に於て x なる所得を有する人數 y を採る代りに，x に對應する y の累加値即ち x 以下の所得を有する人員の合計數を採ることによつて生ずる．x に對應する y の累加値は $\int_{x_0}^{x} y dx$ にて表はされる．度數 polygon の場合には和の形になる譯であるけれども，積分の場合と殆んど同樣に見做され得るから，以下斯樣な場合凡て積分の形で取扱ふこととする．

C 圖



今凡ゆる所得を有する總人數（以下總人數といへばこの意味で用ふることとする）を N とすれば，

$$N = \int_{x_0}^{\infty} y dx.$$

此の總人數を以て $\int_{x_0}^{x} y dx$ を除したものを Y にて表はせば

$$Y = \frac{1}{N}\int_{x_0}^{x} y dx.$$

かくして Y は x 以下の所得を有する人員の總人數に對する比の値となる．Y と x との關係は x, y の關係が B 圖で與へられる場合には C 圖の如くになることは容易に知られる．

次には之と異つた累加形態を考へて，x に對應する xy の累加値，即ち x 以下の所得を有する人員の所得を合計した値をとる．之は前の場合と同樣にして $\int_{x_0}^{x} xy\,dx$ にて表はされる．

今總所得を S とすれば

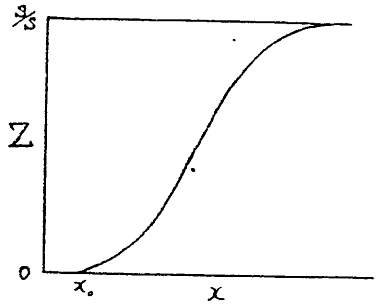
$$S=\int_{x_0}^{\infty} xy\,dx$$

であり，之を以て $\int_{x_0}^{x} xy\,dx$ を除したるものを Z にて表はせば

$$Z=\frac{1}{S}\int_{x_0}^{x} xy\,dx$$

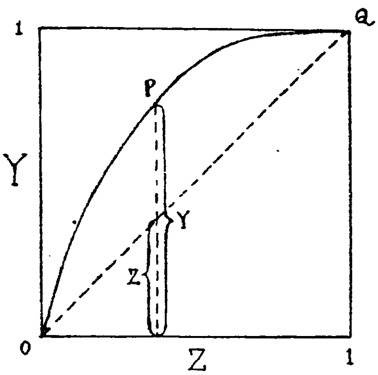
で Z は x 以下の所得を有する人員の所得合計額の總所得額に對する比の値であり，x, Z の關係は D 圖の如くになる．

D 圖



さて此處で次に別に Y と Z とをとり出してその對應關係を觀察する．實際それを圖示して見れば E 圖の如き曲線形を生ずるであらう．此の曲線が實際此の圖の如き形狀を呈し，其の全長を通じて上方に凸なることは次の如くにして證明される．

E 圖



$$\frac{dY}{dx}=\frac{y}{N}$$

$$\frac{dZ}{dx}=\frac{xy}{S}$$

なるを以て，

$$\frac{dY}{dZ}=\frac{S}{Nx}>0 \quad [\because\ x>0\ \ S>0\ \ N>0]$$

又

$$\frac{d^2Y}{dZ^2}=-\frac{S}{Nx^2}\cdot\frac{dx}{dZ}=-\frac{S^2}{Nx^3y}<0 \quad [\because\ y>0].$$

かくして此の E 圖によつて，例へば所得の少い方から數へて總人員中の $100\times Y\%$ の者に屬する所得の和が總所得額の $100\times Z\%$ であるといふ關係，或は恰度その逆の對應關係が直ちに知られるのである．而して此の曲線こそ實に Lorenz curve そのものに外ならない．此の曲線は Y と Z の對應關係を示すものであるから，從つて Lorenz curve は $\int_{x_0}^{x} y\,dx$ と $\int_{x_0}^{x} xy\,dx$, 或は同じことだが，$\sum_{x_0}^{x} y$ と $\sum_{x_0}^{x} xy$ との比率的な關係をあらはすものに他ならない．この兩變量が累加形態であることの外に，それが單純な累加形態でなくして一方が乘積 xy の累加を含む點が特に注意に價すべきことである．此の事實が Lorenz curve をして單なる撒布度でなく，分布の不平等といふ特殊なる概念と關聯するものたらしめてゐると解し得られる[3]

次に Lorenz curve によつて不平等度は如何にして測定せられるかといふに，前述せる如く，それは該曲線がそれに正對する對角線から乖離する程度によつてである．若しも分布が完全平等であるならば，10% の人員が 10% の所得額を享有し，20% の人員が 20% の所得額を受取り，以下同樣であるから，curve は共の對角線と全然一致する．之に反し例へば 90% の人員が總所得額の僅か 10% を占めてゐるといふ樣な場合には (Y, Z) の位置は常に此の對角線上に來らず，むしろそれから著しく離れたところに在るが故に，共の曲線の對角線からの乖離が不平等の指標たり得ることは明らかである．そして曲線の對角線からの乖離をばその全長に亘つて示すものは E 圖に於ける $OPQO$ なる弓形狀の圖形の面積に他ならない．即ち該圖形の面積を λ とすれば，λ の大小によつて原統計値の分布の平等不平等が算定せられる．而して λ が大なるときは不平等度が大，λ が小なるときは不平等度が小なることはいふまでもない．λ の値は最大なるとき $\frac{1}{2}$，最小の時 0 であるが，若し不平等度係數として變域が 0 から 1 なるものが所望せられるとすれば，2λ を以てそれに充てればよいのである．

さて λ は如何にして計算されるかといふに，それは E 圖よりして次の關係が知られるからそれに基いて算定され得る．

$$\lambda = \int_0^1 (Y-Z)\, dZ.$$

實際上は圖を正確に書けば，プラニメターを用ひて λ の値を容易に計測し得るであらう．

此處に一言注意を加ふべきは λ の數値が互に相等しくあつても異つた Lorenz curve が存在し得るといふこと，即ち異つた形狀の Lorenz curve からして相等しい λ の値が得られるといふことである．此のことは實際問題と關聯して相當考慮を要する場合がある．即ち λ は分布の不平等の程度を一應單一の數値に約元して示すけれども，互に內容の異る不平等性をば同一數値を以て示す場合があり得るが故に，只 λ の數値のみに據つて不平等性を論ずるのでは不充分を免れず，曲線の形狀を相互に比較して研究せねばならぬ場合がある．それであるから本來の圖示的な方法が實際問題にとつては相當の意義を有してゐる譯であつて，此の點は具體的な場合に對する適用例から省察さるべきである．

尙此の λ については著しく興味ある性質が發見されて居り，又 λ の値の大きさと原度數分布の形狀又は或る度數常數の値との間に一定の條件の下に於ては一定の關係が存することが推定されるが，それ等は本稿の範圍を超えるから，此處には述べないこととする．

此處に於て次に上に取扱はれた樣な財產所得等の各個人への分配現象と異る場合，即ち人的集中現象以外のものについての Lorenz curve の適用の如何を論ぜんとするが，その際この人的集中以外の場合の中の代表的なものである地理的集中現象をとつて，それについて考察する．其他の集中現象はその性質が若干不明瞭であり，又著しく他と異る樣にも見られるが，方法的にはむしろ此の地理的集中現象の場合の適用の仕方の擴張が可能と考へられる．尤もこの點は尙保留しておきたい．

さて地理的集中現象であるが，それは例へば人口が國土の全範圍に分布するに，その全範圍に平等に分布せず，一部の範圍には多い量が，他の部分には少い量が配分されてゐるといふ樣な場合である．それで此の樣な場合の分布の集中乃至不平等の程度を如何にして分析するかといふのである．

かゝる場合に若し其の上に人口の分布が觀察されるところの國土なるものが全然不可分のもので，一切の分割觀察が拒否されるならば，凡そその分布の性狀を量的に表現することは不可能である．かゝる量的表現が可能なるためには國土の全地域が幾つかの成る可く多數の部分に分割され，その個々の部分が別々に觀察され得なければならない．然る後にこの各個の部分の觀察の結果が綜合されて此處に國土全體としての分布の狀態が云爲されるのである．それ故に此處で必然的に要求される手續は，全領域をば幾つかの部分に割つて觀察すること，即ち各部分地域の夫々が保有する人口量を調べることである．そしてかやうに部分に分割して觀察する場合には，其處にどうしても地理的單位なるものが必要とされ，その各地理的單位が夫々保有してゐる人口量が明かにされねばならない．その結果に基いて始めて分布の平等不平等が取扱はれ得るのである．

惟ふに此の際若しも此の地理的單位なるものが，國土面積を若干個に全く等分する樣な互に齊一なものであるならば，かゝる地理的單位をば前記人的集中現象に於ける人員と全く同じに取扱つて共の個數を人口量なる變量に對應する度數の如く考へれば，それによつて Lorenz curve の適用の問題も全く何の困難もなく解決することが出來るであらう．ところが少くとも社會統計に關する限り，多くの場合といふよりも殆んど凡ての場合に地理的單位としては左樣な各個齊一なものが現實に得られず，それは例へば國家の行政區劃といつた樣なもので，一つ一つがその大きさを異にしたものである．それであるから，それに對しては上記の如き取扱ひはなし得ない譯であつて，若し敢てそれを各個の單位が齊一なものの如く取扱つたならば，それはむしろ非常に誤つた結果をもたらすことになる．

其處で斯樣な場合は吾人にとつて如何なることを意味するかを再檢討して見るに，各地理的單位が夫々固有の人口數を保有するが，その個々の人口數は夫々異つた面積內に保有せられてゐるのであり，從つて夫々異つた基礎に立つものであるから，此等の人口數そのものは互に比較されても意味がない譯である．それでは此の場合互に比較性を有する量は何であるかといへば，それは各地理的單位の保有する人口の絕對數でなくして，それを夫々の地理的單位の面積にて除したるもの即ち各地理的單位の人口密度である．而して此の人口密度の數値は各地理的單位に固有のものであるから，其處にその各人口密度の差異，從つて密度の分布を考へ得る．その際異る密度は夫々の地理的單位の大きさの範圍內に於てのみ妥當するのである．此の樣な關係は一體 Lorenz curve により如何に表示されるであらうか．

設例としては引續き人口の地理的分布をとつて考察する．人口數を x，地理的面積を y にてあらはし，更に地理的單位を例へば府縣としよう．然らば全國各府縣についてその人口と面積を調査することによつて相對應する x, y の値の組が多數（此處では府縣の總數だけ）得られる譯である．けれども此の場合に直接 x と y との間に度數分布の如き關係を表示することは一般に不可能である．何となれば人口が x と $x+\Delta x$ との間にある樣な土地面積といふものは此の場合には抑も意味をなさないからである．然し今地理的單位の，即ち此處では各府縣の，人口密度を h にてあらはすこととすれば h と y との間には度數分布の如き關係が存在すること疑ない．h に或る範圍の値を與へた場合に，かゝる人口密度を有する土地面積が幾何なりやは所與のデイタより當然導き得られる．

それであるから此處で次に h と y とから Lorenz curve の基本をなす如き累加形態を構成せんとし次の如き大きさを導く．

$$N=\int_{h_0}^{\infty} y\,dh \quad , \qquad Y=\frac{1}{N}\int_{h_0}^{h} y\,dh$$

$$S=\int_{h_0}^{\infty} hy\,dh \quad , \qquad X=\frac{1}{S}\int_{h_0}^{h} hy\,dh \quad .$$

此處に N は國土總面積，Y は人口密度が h 以下の部分の面積の和の國土總面積に對する比の値であることが知られる．

又定義により $h=\dfrac{x}{y}$ なるを以て之を上の S 及 X の式に代入すれば，

$$S=\int_{h_0}^{\infty} x\,dh \qquad X=\frac{1}{S}\int_{h_0}^{h} x\,dh$$

となり，S は總人口を，X は人口密度が h 以下の部分の人口の和の總人口に對する比の値なることが判明する．

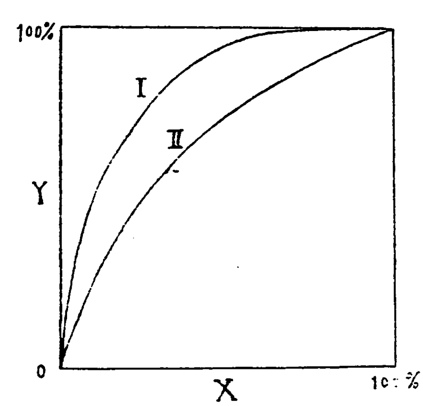
從つて此の場合には Y と X との關係，從つて $\sum_{h_0}^{h} y$ と $\sum_{h_0}^{h} hy$ との比率的對應關係を示すところの Lorenz curve は結局に於て單に $\sum_{h_0}^{h} y$ と $\sum_{h_0}^{h} x$ との比率的對應關係を示すものに歸着することが窺ひ知られる．而して Lorenz curve による不平等度の測度たる前記 λ は此の場合には次式を以て與へられる．

$$\lambda=\int_0^1 (Y-X)\,dX \quad .$$

而して此の場合にも $0\leqq 2\lambda\leqq 1$ である．

要するに此の樣な場合には一般に x と y との直接の對應關係は示され得ないけれども，その單純な累加形態 Σx と Σy との關係は Lorenz curve の形態に於て却つて直接示され得るといふ興味ある事實が見出される．かゝる地理的集中現象への同曲線適用の仕方は他の類似の場合にも及ぼすことが可能であらう．又地理的集中現象にあつても，必ずしも地理的面積が基準とされずともよく，例へば一定現象の人口に對する地理的分布といふ樣なものを觀察する際には，地理的人口が面積に代つて基準とされ，同樣に取扱はれ得るであらう．[5]（F 圖は昭和十二年末本邦の工場數の分布を示す曲線で，I 曲線は府縣面積基準，II 曲線は府縣人口基準である．）

F 圖



最後に吾人は Lorenz curve の應用例を今一つ追加しよう．それは一般に變化率を示す種々なる數値の分布に對し，之を適用して解析し得る點を指摘することである．

今 t_0 時點に於て A なる大きさを有する集團が t_1 時點に於て B なる大きさになつたとする．而して $B-A=C$ である．即ち該集團は t_0 から t_1 迄の間に C だけ增大し，其の增加率は $\dfrac{C}{A}$ である．ところで之は其の集團全體について該當することであるけれども，若し此の集團をば幾つかの部分集團に分つて觀察するならば，其の各個の部分集團に對しては恐らく決して一樣に $\dfrac{C}{A}$ なる增加率が適用

されるものではなく，各部分集團毎に様々の異る増加率（或ひは減少率）を有してゐるであらう．今此の各部分集團の増加率（或ひは減少率）を $\frac{c}{a}$ にてあらはす．即ち各部分集團の t_0 時點における大きさを a とし，t_1 時點に於ける大きさを b とし，$b-a=c$ とする．さうして $\frac{c}{a}=v$ とすれば，v の種々なる値に對應する a が素材から獲られる理である．v に或る範圍の値を段階的に興ふるならば，v に對する a の度數分布が描かれる．

此處に於て前の場合と同様にして $\frac{1}{A}\int_{v_0}^{v} adv$ と $\int_{v_0}^{v} av\,dv \Big/ \int_{v_0}^{\infty} avdv$ との對應を示すところの Lorenz curve が生ずる筈である．然るに $v=\frac{c}{a}$ なるを以て

$$\int_{v_0}^{v} av\,dv = \int_{v_0}^{v} cdv \qquad \text{及び} \qquad \int_{v_0}^{\infty} av\,dv = \int_{v_0}^{\infty} cdv = C \ .$$

從つてそれは $\sum_{v_0}^{v} a$ と $\sum_{v_0}^{v} c$ との對應を示すものである，此の場合 c に負のものが存在すれば，今迄の場合と異つて曲線は第二象限の方へはみ出る様な形となる．然しこの場合も λ は前と同様にして算出し得る．

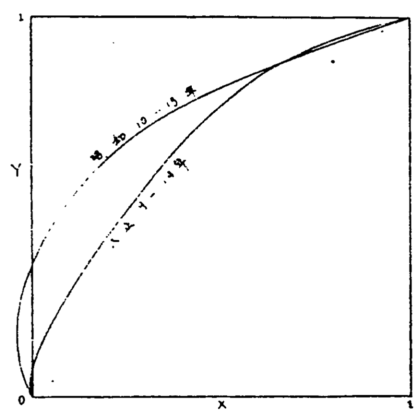
$$\int_{v_0}^{v} adv = Y \ , \qquad \int_{v_0}^{v} cdv = X$$

とおけば

$$\lambda = \int_0^1 (Y-X)\,dX \ .$$

但しこの場合 2λ の値は 1 を超えることがある．尤もそれは不平等が非常に著しい場合のみであるがその點から云ふも，此の場合は曲線はむしろ圖示的なものとしての役割が大であるといへよう．

G 圖



玆に本邦の人口増加率の府縣別差異を此の方法で描いた試みを揭げる．之は大正 9 年 ～ 14 年の 5 ケ年間の變化と昭和10年～15年の 5 ケ年間の變化とを對照して示したものであるが，此處に描かれた曲線によつて，此の二つの時期の人口増加率の內容の著しい相違が明瞭に觀取される．大正年代の人口増加率の府縣別差異は夫れ程著しくないが，昭和年代のそれは非常に著しくなつて　謂はば甚だしく不健康な様相を呈し來つたことは，全く不用意に看過し難い重大問題といはねばならぬ．

之を要するに Lorenz curve は，先に明らかにした如き度數分布の累加形態に基くものであるが，その適用は富や所得の分布に關する人的集中の問題のみに局限されることなく，其の他にも廣汎な應用領域が見出され，而も相當興味ある結果が獲られることは注目に價することといへよう．

1) Kurtz and Edgerton, Statistical Dictionary. New York. 1939.
2) 田村市郎，分配狀態の測定法に就て，日本統計學會年報，第 7 年（昭和13年），p. 10.
3) 田村市郎，前　揭，p. 2.　同氏が實體分布と呼ぶものは惟うに xy-product を含めるものである．
4) 米澤治文，工業分布統計への一寄與，日本統計學會年報，第10年（昭和16年），p. 19-20.
5) 米澤治文，工業の地方的分布狀態分析の諸方法に就て，東北帝大研究年報「經濟學」第12號（昭和15年），p. 100.