慶応義塾大学講座経済学

寺尾琢磨・著

慶応出版社

第１４

昭和１６

寺尾琢磨著

統計學

14

慶應出版社刊行

目次

第一章　緒論

第一節　統計學とは何か

人智の撓みなき進歩は日に日に未知の世界を開拓して、曾ては玄妙不可思議とされてゐた幾多の謎を解決し、また絶えず新たなるものを創造して、知識の範圍を擴大し、以て生活の内容を豐富にしつゝある。併し他方これが爲に吾人の生活樣式が複雜となり、事物の本質も相互の關係も、これを知る事益々困難を加へるに至つた事も亦爭へないのである。文化の進展が本質的に人類の幸福を增進せしめるか否かは極めて疑問なのであつて、寧ろ簡素な原始の世界こそ、或ひは吾人の求める桃源郷であるかも知れない。併し乍ら翻つて惟ふに、總て學問なるものは、與へられた事實を事實として有りの儘に觀察し、これを合理的に説明するのが本領である以上、その道義的批判の如きは次の階段に讓らねばならぬ。嚴格な科學的見地に立つ限り、吾人の第一の課題は、如何にすれば事實を最もよく觀察し最もよく説明し得るかといふ事でなければならぬ。

事實を觀察し説明する方法、卽ち科學的研究法は、事實の種類に從つて自ら異らねばならぬ。事實はそれを發生せ

しむる原因に從つて二つの範疇に分ける事が出來る。即ち若しそれが自然的物理的原因から生じたものであれば、自然的事實又は自然現象といひ、これに反して人間の意思及び行爲から生じたものであれば、社會的事實又は社會現象といふのである。この二つの事實は發顯の様式に大なる相違がある。蓋し自然現象はその原因の性質上、極めて機械的であつて、所謂自然法則なる嚴格な法則によつて律せられてゐるに對し、社會現象は元來自由な人間行爲の集積に外ならぬから、多分に偶然性を帶び、果してこれを律する何等かの法則あるや否やも甚だ疑問なのである。人間は著しい程度に感情の動物であるから、その集積たる社會現象は決して一律の原則を以て説明されるが如きものではない。「多來らば春遠からじ」とか、「生ある者は必ず死す」とかの自然的命題ならば疑問の餘地はないが、反之、人口增加は生活程度を低下せしめるといふ様な社會的命題は決して一義的に肯定出來るものでない。低下せしめるときもあるし、反對に向上せしめるときもある。それは其時の社會狀態によつて決定されるのであつて、斯かる社會狀態は結局は人間行爲の產物に外ならないから、要するに社會現象は自然的產物ではなくて歷史的經驗的產物と言はねばならぬ。

斯く社會現象は自然現象と異つて本質的に箇別的なものであるが、然らばこれを律する何等の法則も無いかと言へば、決してそうではない。これは二つの理由から説明出來ると思ふ。第一に人間の生活は絶えず自然の制約下にあること、第二に人間は大體に於て合理的な動物であり、且つ合理的ならしめんとする法律・制度・道德等の制約下にあること之である。卑近な例を考へて見よう。三伏の候に敢へて厚着することは自由である。が、餘程の事情が無い限り、夏は薄着するのが普通である。これは自然制約の結果に外ならない。二人の妻を持つ事は不可能ではないが、併しこれは重婚罪を構成し法律の處罰を免れないから一般に行はれない。又商品價格を倍にする事は勝手であるが、價格統制法や自由競爭によつて抑壓されて了ふ。これは社會的制約の結果に外ならない。吾人は絶えずこの二つの制約下に在るが故に、社會生活は自ら或る軌道に沿つて進行するのであつて、社會現象にもこれを律する法則があるといふのは蓋しこの意味に於てゞある。勿論この社會法則は自然法則の持つ普遍妥當性には缺けてゐる。自然法則の先驗的なるに對し、社會法則は社會的集團から歸結されるに過ぎず、常に時と所とによつて異同あるものである。即ち法則と言はんよりは寧ろ傾向或ひは規則性と呼ぶ可きであらう。

然らば社會科學の目的は社會的事實を觀察し、その規則性を把束するに在る事は自ら明かであらう。そしてその際とるべき方法が自然科學のそれと異らねばならぬ理由は、社會的事實なるものは單に經驗によつてしか確定されないといふ事である。吾人の經驗には言ふ迄もなく質的なものと量的なものとがある。私がこれから述べようとする統計學なるものは、實に社會現象の量的經驗を取扱ふ事によつて社會的規則性の確定に寄與せんとする學問なのである。

斯く統計學の對象は社會的現象だといふ意味は、個體に非ずして集團を對象とするといふ事である。集團は素より個體から構成されるから、個體を無視して集團を想定する事は不可能である。併し社會的集團は個體の無機的結合ではないから、自ら個體とは別箇の性質を示す事になる。この性質が廣範圍に亙つて認識されるならば、これを以て社會法則又は社會的規則性と認めるのである。然らば如何にして個體を集團に[illegible]め、集團からその性質を抽出し得るか。こゝに統計學の課題を發見し得るのである。

統計學の本質に就ては專門家の間に於てすら議論のあるところであつて、後に述ぶるが如く、實質的社會科學と解するものと統計的研究法の學即ち形式的方法學と解するものとある。私は後者、即ち統計學を以て一箇の方法學と解

するものであるが、併しこの派にも社會現象の方法學に限定するものと、一般集團現象の方法學と解するものとが對立してゐる。社會現象に限定する事は、既述の社會現象と自然現象との根本的相違からして、最も異論の餘地が少いと思はれる。併し自然的集團現象を全く除外するのは聊か早計であらう。蓋し自然現象は物理的因果的關係によつて規定されるものではあるが、この關係が悉く吾人の知るところとは言へないのであつて、幾多の自然現象は、恰も社會現象と同じく、偶發的外見を呈してゐるのである。天體・氣象等の現象には特にこの感が深い。斯かる場合には、集團の規則性を抽出するを目的とする統計的方法は極めて有力な武器となる筈である。即ち統計的方法が自然科學の探究に適用される事は毫も不合理ではない。併し自然法則と社會法則との根本的差異に基き、統計的方法の本來の使命が、依然として社會科學の領域に在る事には變りがないのである。

私はこの講義を特に經濟學研究者の爲に執筆する。上述の如く私は統計學を以て統計的方法の學と解するものであるから、必ずしも經濟的集團を取扱ふ必要はないが、併し本講座が經濟學の學修を目的とする事と、並びに經濟的集團はそれ自身特殊の性質を持つ事からして、出來得る限り經濟現象への適用を中心としたいのである。尤も純粹の應用については、別に與へられた「經濟統計學」なる講義に於て詳しく述べるつもりである。尙ほ方法學としての統計學は、元來確率理論に立脚する所謂數理統計學を中核とするのであるが、讀者の大半に高等數學の知識を豫定し得ない以上、努めて數理的說明を避けねばならぬ。缺く可からざる場合には、脚註の形式によつて補つて行きたい。

第二節　統計學史の概要

或る科學の課題・使命又は本質を理解する爲には、一應その歷史的發展の跡を辿つて見る必要がある。蓋し科學も亦一種の歷史的產物であつて、時代によつてその意義は絕えず變化しつゝあるからである。そしてこの意味に於て統計學ほど時代と共に根本的改革を經驗したものは稀であつて、從つて統計學なる同一名稱の下に理解される知識體系は決して單一ではないのである。濫觴時代の統計學と現今のそれとは全く別箇の趣きを呈し、それが同一名稱を以て呼ばれる事すら甚だ奇異の感を與へる程である。斯かる時代による變化は畢竟缺陷が次第に訂正されるからで、從つて發展の經過を顧ることは前車の轍を避けしむる最善の方法と考へられるのである。

統計學の歷史は統計調査の歷史と混同されてはならぬ。土地・人口・資源等の調査は多少とも組織的な國家を形成された遠い昔から行はれてゐる。これは徵稅・賦役等の必要から當然の事で、例へば埃及時代、ピラミット建設の爲に人口調査の行はれた記錄があるし、支那の禹帝、我國の崇神天皇の御代にも既に人口調査が實施されたといふ。玆では斯かる調査には觸れる事なく、學問としての統計學が如何にして生れ、如何なる經路を辿つて今日に至つたかを略述しよう。惟ふに統計學は國勢學派・政治學派及び確率理論なる全く異る三つの源泉から出發したもので、それが前世紀中葉に一應綜合され、更にその後社會統計學派と方法論學派の二箇に分裂して相對峙してゐるのである。

（一）國勢學派（大學統計學派）

國家の成立と共に、國家事務遂行の必要上、國土や住民等について各種の統計調査の行はれるに至つた事は右に述べたが、斯かる調査によつて資料が豐富となれば、それが次第に組織的知識體系に、即ち一箇の學問に、編成されるのは當然であらう。統計學（即ち統計に關する組織的知識體系）が國狀記述の學として誕生したのは蓋し最も自然の

趨勢である。從つてこの意味に於ける統計學の淵源は極めて古く、既に希臘時代に遡る事が出來るのである。碩學アリストートル (Aristotle; 384—322 B. C.) の「政治學」なる大著は、その後散佚して今日は僅かにその片鱗を留むるのみであるが、例へばアテネの政治機構に關する記述の如きは、明かにこの傾向の先驅と見られる。社會及び國家の組織が更に發展擴大した羅馬時代には國勢調査すら行はれた記錄があり、國勢學としての統計學の存在した事は疑ふべくもない。羅馬帝國滅亡に續く永い暗黑時代は、人間進步の歷史に於て、單に白紙と言ふよりは寧ろ逆轉を意味し、人類の知的精神的活動は全く中斷されてゐたのである。併し永い闇の世界は、十五世紀に至り伊太利に勃興した文藝復興運動によつて、新たなる光明の世界に移り、近代文明の端緒は開かれた。中絶せる國家記述の學も玆に再び擡頭し、サンソヴィーノ (Francesco Sansovino, 1521—86) 及びボテーロ (Giovanni Botero, 1540—1617) がその中心となつたのである。

文藝復興の波はアルプスを越えて全歐に漲つて行つた。この波に乘じて各地に輩出した幾多の統計學者のうち、特に注目に値するのは佛蘭西のパキエ (Etienne Pasquier) 及びダヴィティ (Pierre d'Avity) であらう。パキエの「佛蘭西研究」(Recherches de la France, 1581) は佛蘭西帝國成立の歷史を論じ、社會機構を特に法的觀點から明かならしめんとした不朽の名著である。ダヴィティの「國家及び君主」(Les Etats, Empires, Royaumes, Seigneuries, Duchez et Principautez du Monde, 1614) は過去現在の風俗習慣や、君主の系圖・國家の收入・通貨・宗教・軍事等を仔細に記述したものである。

併し乍ら國勢記述の學としての統計學が確たる地位を獲得したのは獨逸に於てゞある。主として地理學的記述を行つたフランク (Sebastian Franck) 及びミュンスター (Sebastian Münster) は別としても、ゼッケンドルフの「獨逸君主國」(Seckendorff, Deutscher Fürstenstaat, 1656) の如きは、貴族に行政上必要な知識を授くるを目的としたもので、獨逸統計學の發端を爲すものである。併し、獨逸統計學が特に國勢學 (Staatenkunde) 又は大學統計學 (Universitätsstatistik) なる特殊の學派を形成するに至つたのは、ヘルマン・コンリング (Hermann Conring, 1606—81) 以後の事に屬する。彼はヘルムシュタット大學教授として各國情勢の比較を講義し (その中には日本に關するものもある)、有能なる幾多の後繼者を得た。彼は物的・目的的・形式的及び手段的の四原則に從つて記述對象を分類した。物的とは國土と人口、目的的とは國家の目的、形式的とは機構及び行政、手段的とは財政及び軍備を指すのである。彼の記述は數字に據らず、單に例へば某國は人口稠密なりとか稀薄なりとか言ふ形式を採つた。この言語式表現は、國勢學派の特徵であり、政治算術又は今日の統計學と根本的に異る點である。

統計學を意味する Statistik なる文字を始めて使用したのは、コンリングの門弟アッヘンワル (Gottfried Achenwall, 1719—72) である。國勢學としての統計學は彼によつて大成されたと言つてよく、統計學史上逸す可からざる學者である。彼がゲッチンゲン大學教授であつた關係上、彼及びその一派をゲッチンゲン學派と呼ぶ。代表的後繼者としてはシュレーツェル (Schlözer, 1735—1809) を擧ぐべきであらう。

國勢學派が言語的表現に終始した主たる理由は、當時の專制君主國に於ては、假令統計調査は行はれても、その結果の公表を見る事は無かつた事である。而も彼等は自己の方法に絕大な自負を抱いてゐたため、和蘭のアンヘルセン (Anchersen) なる學者が數字による表示統計 (Tabellenstatistik) なる新表現方法を提唱するや、「表の奴隷」と罵

つて、爲に一八〇六年前後、激しい論爭を惹起した。併し英國の政治算術派の影響が次第に侵入し來ると共に、國勢派の方法にも漸次變化を來し、數字的記述が採り入れられるに至つた。この傾向の代表者としてはビュシング（Büsching, 1724—93)、クローメ（Crome, 1753—1833）を擧げて良からう。

一八五〇年クニース（Knies）が「獨立科學としての統計學」なる劃期的名著に於て統計學なる名辭の混亂を指摘し、政治算術がこの名稱に値すると說くに及んで、統計學としての國勢學は終りを告げた。自己に命名した名稱が自己に値せざるものとして他に剝奪されるに至つたのは、大なる皮肉と言はねばなるまい。そして社會の複雜化と科學の飛躍的進步は、國勢といふが如き廣汎な領域を一科學の對象とするを許さゞるに至り、結局國勢學そのものも、歷史學・地理學・政治學・社會學・經濟學等の幾多部門に分解される事によつて、姿を消して了つたのである。

（二）政治算術（Political Arithmetic）

獨逸に於てコンリングが國勢學の建設に努力してゐた頃、英國には政治算術なる別趣の統計學が誕生した。一六六二年ロンドンの一商人ジョン・グラウント（John Graunt, 1620—74）の著した「死亡表に關する自然的政治的觀察」即ちこれである。當時の英國は頻々たる流行病、急激なる都市の膨張によつて一般衞生狀態が著しく低下し、當局は銳意結婚・出生及び死亡の記錄を整備するに努めた。一六〇三年以來これらに關する記錄は每週公表され、年末にはその合計が發表された。即ち大陸諸國と異り、英國には數字的材料が豐富に備つてゐたのである。グラウントがこの材料に着目したのは極めて賢明であつた。蓋し彼によつて未開の領域が開拓され、豐富な收穫が約束されるに至つたからである。

彼はその著書の表題の示すが如く、主として死亡統計に關する硏究を中心としたが、それに附隨して出生・婚姻・住民數測定等に觸れてゐる。その資料の取扱方、計算方法乃至は結果の解釋等については、今日から見れば遺憾の點の多い事は勿論であるが、併し一見亂雜極まる人口現象も、大量觀察によれば驚く可き規則性に律せられてゐる事を發見したのは、彼の不朽の功績と言はねばならぬ。

彼の硏究は直ちに友人ペティ（William Petty）の繼承するところとなつた。政治算術なる名稱は實にペティの著書（Political Arithmetick, 1690）に基くのである。彼は人口數・人口倍加期間等につき幾多興味ある所說を吐いてゐる。併し例へば每年人口百二十人につき一人の割で增加すれば、百二十年にして人口は倍加すると言ふが如き、複利の原則を忘れた計算を行つてゐるのは、數學的知識の乏しさを如實に曝露したもので、彼の如き一流の學者にして猶ほ斯くの如しとすれば、一般の知識水準が如何なるものであつたかは容易に想像し得よう。

グラウント及びペティに較べて遙かに科學的才能を示したのはハレー（Edm. Halley, 1656—1742）である。當時獨逸のブレスラウ市は久しきに亘る出生・死亡統計を備へてゐたが、偶々當時の世人の間に、一般に人間には七年を周期とする危險期があり、特に四十九歲と六十三歲が最大の厄年だとの信念が行はるゝを見て、ノイマン（Kaspar Neumann）なる人が右統計によつて、その眞僞を明かならしめんとした。その報告は當時の大哲學者ライプニッツ（Leibnitz）に提出され、後者は更にこれをロンドンの王立協會に移送した。ハレーはその囑を受けて右の仕事を繼續し、有名な“An Estimate of the Degrees of the Mortality of Mankind drawn from curious Tables of the Births and Funerals at the City of Breslaw, 1693”に纏め上げたのである。統計學史の大家ウェスターガー

ド教授が「生命統計に對するハレーの貢獻は如何に高く評價しても足りない。驚嘆すべき思考の明白さを以て、彼は自己の問題に正しき形を與へた。不幸にして彼の後繼者は常に必ずしも彼を理解せず、爲に彼の研究も、當然値すべき榮譽を逸して了つた」と述べてゐるところから、その學問的價値は、容易に判斷されるであらう。生命保險の基礎たる生命表なるものは、實にハレーの勞作に負ふのである。

英國に興つた政治算術は間もなく大陸に傳はり、最後に十八世紀後半に至り獨逸のジュースミルヒによつて完成されたのであるが、その永い中間の期間に輩出した代表的人物はヴォーバン (Vauban)、デヴナント (Devenant)、キング (G. King)、ベルヒ (Anders Berch)、スマート (J. Smart)、モアーヴル (de Moivre)、エルヴウス・(Pehr Elvius)、ワルゲンティン (Per Wargentin) 等である。就中、最後に擧げたワルゲンティンは、瑞典屈指の天文學者であり、而も人口統計の一大恩人である。著名なマルサスの人口論の中に頻りに引用され、人口學者に取つては特に有名となつてゐる。併しこれらはいま一切省略し、最も重要なジュースミルヒに移らう。

プロシャの僧侶ジュースミルヒ (J. P. Süssmilch, 1707—67) は十八世紀の幾多學者と同樣、人口問題を以て政治の中心と考へ、人口增殖に關する殆ど凡ゆる問題に關心を寄せるに至つた。彼をして人口を斯くも重視せしむるに至つたものは、實に彼が聖書のうちの「産めよ、殖えよ、地に盈てよ」なる一句を以て、神が人間に下した至高の命令と考へたからに外ならぬ。彼の大著「人類の出生・死亡及び增殖から證明さるゝ、人類變化に於ける神の攝理」(Die göttliche Ordnung in den Veränderung des menschlichen Geschlechts, aus der Geburt, dem Tode und der Fortpflanzung desselben erwiesen, 1741) は、表題そのものが既に宗教的色彩を濃厚に示してゐる。洵に彼の意圖は、人口增殖の原理のうちに人智の及ばざる深き神の攝理の示顯さるゝを說き、これによつて神の廣大無邊なる仁愛を頌へんと欲したに外ならぬ。從つて動機と共にその解釋が非科學的なことは止むを得なかつたが、併し研究の方法に至つては、今日と雖も首肯されるものが甚だ多いのである。彼はグラウント、ペティ等の方法を巧みにプロシャ諸州の新たなる資料に適用し、人口現象に於ける驚嘆すべき幾多の規則性を發見したのであつて、例へば女兒二十人につき男兒は二十一人生れること、而も男兒の死亡率は女兒よりも高く、從つて結婚年齡に於ては男女數は等しくなる傾ある事を說いてゐる。これは今日に於ても略ゝ眞實であるが、彼は右を以て一夫一妻主義を實現せんとする神の意圖に出づるものと解釋したのである。併し彼が都會と農村の死亡率を比較し、都會の不健康を指摘して自然的生活の意義を說くが如き、後の都市政策に重大な暗示を與へたものである。彼を以て近世統計學の開祖と仰ぐ人の尠くないのも强ち不合理とは言へないと思ふ。

（三）確　率　論

事象出現の確からしさに關する理論を確率論といふ。元來數學的哲學的理論に外ならぬが、併し近代統計學は確率論と切離して考へる事は不可能であるから、玆に簡單に如何にしてこの理論が成立し如何にして統計學と結びつくに至つたかを說明して置かう。

中世の荒唐無稽な鍊金術が近代化學の母胎となつたり、印度へ到達する目的で西航したコロンブスが圖らずも米大陸を發見したり、人間の努力は屢ゝ思はざる結果に到着するものであるが、この事は確率論の發展に於ても亦同じである。卽ちこの理論は元來賭博者が平常の經驗に基いて當率を算出した事から始まつたのであつて、十六世紀の伊太

利の著名な賭博者カルダノ（Cardano）がこれに關する最初の理論を樹てたと言はれてゐる。この問題は忽ち數學者の興味を惹き、パスカル（Pascal）、フェルマ（Fermat）等によつて次第に一箇の理論體系に高められるに至つたが、特に夫のジャコブ・ベルヌーキ（Jacob Bernoulli, 1654—1705）及びその甥ダニエル・ベルヌーキ（Daniel Bernoulli, 1700—1782）によつて著しい進歩を遂げたのである。死後に出版された前者の「豫望論」（Ars Conjectandi, 1713）は、未定稿ではあるが、確率論に於ける劃期的著作である。該書は四部より成り、（一）に於ては當時これに關する最良の著作と言はれたユジャン（Chr. Huygens）の「賭博論」の解説を取扱ひ、（二）は順列及び組合せを論じ、（三）に於ては賭博の各種問題に觸れ、最後に（四）に於て確率論の道徳及び經濟への適用を論じてゐる。この最後の部分は未完成ではあるが、「觀察度數が非常に大なれば、經驗的確率は先驗的確率と略〻一致する」といふ所謂ベルヌーキの定理が展開されてゐるのである。この定理こそ、統計的研究法の理論的根據、即ち大數法則の原理に外ならないのであつて、今日の統計學が確率論と切離して存在し得ない理由はこゝに在るのである。

ベルヌーキの定理は何故直ちに統計學に採り入れられなかつたか。惟ふに第一には統計學者が數學的素養を持たなかつた事、第二には斯かる數學者が餘りに抽象的な議論に馳り過ぎて實際への適用を困難ならしめたことを擧げねばなるまい。併しその後ラプラス（Laplace）、ラグランジ（Lagrange）、ガウス（Gauss）等の巨匠によつて、この理論が飛躍的進歩を遂げ、數字的觀察の結果に對する信頼度・最小自乘法・二項定理等の主要理論が展開さるゝに至り、數量に對する統計學者の態度も根本から改められる事になつた。この業績は十八世紀後半の偉大な統計學者ケトレー並びにその後繼者達の努力に負ふのである。

（四）ケトレーの近代的統計學

既に述べた通り、統計學としての國勢學は十九世紀に入つてからは勢力を失墜し、政治算術のみが獨りこの名稱を擅にしてゐたが、併しその內容は人口統計の域を脱せず、その方法に至つては極めて素朴なるを免れなかつた。他方數學は著しい進歩を示したとは言へ、その適用は殆ど自然科學、就中天文學に限定されてゐたのである。統計學の對象を單なる人口集團から一般社會集團に、否、自然的集團にすらも擴大すると共に、その方法に數學、就中確率論を導入して新たなる原野を開拓したのは、白耳義人ケトレーの不朽の功績である。彼を目して近代統計學の父と言ふのは決して過言ではない。固より今日の統計學はケトレーの建設し又は建設せんと企てたものとは著しい懸隔があるけれども、それは後に說明する通り、ケトレーの行過ぎを矯正した結果に外ならないのである。

彼は最初數學と天文學とを專攻し、測地學や物理學にも興味を抱いた。パリーに遊んでラプラス、ポアソン、フンボルト、フーリェ等と交誼を結び、特にフーリェに就て數學を、ラプラスに就て確率理論を究めたが、次第に彼の興味は自然科學から社會科學に移つて行つた。統計學者としての彼の活動は極めて廣範圍に至り、一八二九年の國勢調査の計畫、ロンドン統計協會の設立、統計中央委員會、國際統計會議等何れも彼の業績である。又彼によつて官廳統計の改善、統計方法の進歩は著しく促進されたのである。併しこゝでは斯かる實際的及び技術的問題を割愛し、彼の主著たる「社會物理學、或ひは人間能力の進歩に關する一論」（Physique Sociale, ou Essai sur le Développement des Facultés de l'Homme）を一瞥しよう。本書に從へば、社會現象は一見恰も放恣な人間自由意思の所産に過ぎず、何等これを律する法則なきが如くであるが、而も統計的研究の結果は全くこの豫想を裏切り、自然現象と同じく、

嚴格なる自然的物理的法則の支配下に在るのである。人間の生理的道德的性質は各種の攪亂的要素によつて相互に異るけれど、社會全體を一括して觀察するならば、大數法則及び誤差法則（これらは數學的確率論の援用である）によつて、これらの相違は相殺され、こゝに一定の典型的「平均人」なるものが出現する。即ち平均人は一切の攪亂的要素から解放された本質的人間であるから、その性質を規定するものは普遍妥當性を有する自然法則に外ならない。斯くて平均人の諸性質を明かにする事が、社會科學をして自然科學と同一の精密科學に引上げる唯一の方法となる。然るに平均人なるものは畢竟社會全體の人間の平均によつて知り得るのであるから、統計的方法こそ、それに達する唯一の手段たるものである。

これが彼の根本思想であつて、その「社會物理學」が凡ゆる人間の肉體的精神的諸性質を一々とり上げてゐるのはこれが爲である。そして肉體的性質に關するものは主として人口統計であり、精神的性質に關するものは道德統計であるから、彼の主力は全くこの二方面に傾注されたのである。

斯くの如くケトレーは社會現象を自然現象と同一視する結果、人間の自由意思は活動の餘地が無い事になる。彼の後繼者が宿命論的見解を採るに至つたのも不思議はない。その代表的學者としてヘルシェル（J. Herschel）及びバックル（T. Buckle）を擧げてよからう。ヘルシェルに從へば、一見人間の自由意思と見ゆるものは、實は環境に對する反作用に過ぎない。換言すれば人間は全く環境の產物なのである。バックルはケトレーの思想を歴史學に適用し、自然科學的方法を提唱したので有名である。彼の大著「英國文明史」はこの意味に於て統計學史上の一大金字塔でもある。併し乍らこの社會的宿命論は一般の歴史家、又は神學者達の憤激を招き、暫くの間は學界に大なる波紋を捲き起した。併し後に至つて獨逸歴史學派の巨匠シュモラー又はワグナー等が統計的法則性の妥當限界を確定するに及んで、いつしか姿を消して了つたのである。

要するにケトレーの根本的誤謬は、社會的規則性を尊重する餘り、社會科學と自然科學との區別を無視するに至つた點に在る。惟ふにグラウントによつて統計的研究法の導入されて以來、殆ど全部の統計學者は、社會現象の規則性なる新發見に眩惑され、かゝる規則性からの乖離とか、乃至は規則性そのものゝ變動性とかについては、殆ど考へが及ばなかつた。このクライマックスがケトレーの「社會物理學」なのであつて、彼以後の統計學者に取つては、斯かる統計學萬能の思想を正しい軌道に引き戻すことが最大の課題となつたのである。

（五）現代の統計學

社會統計學派
方法論派

ケトレー派に對する修正は二箇の方向に於て行はれた。一を社會統計學派といひ、他を方法論派といふ。その中心地に從ひ、前者を獨逸學派、後者を英米學派といふ事もある。社會統計學派に從へば、社會現象は根本的に自然現象と異るものである。後者が普遍妥當の物理法則に律せらるゝに反し、前者は全體として見れば或る一定の傾向乃至規則性を示すといふだけであつて、決して先驗的にこれを律する嚴格な法則の如きはあり得ない。社會はケトレーの言ふが如き抽象的な平均人の集積ではない。即ち統計學の目的は決してケトレーの努力した自然的法則の發見に在るのではなく、單に右述の如き傾向乃至規則性の探究に在るのである。この目的を達する爲には、社會現象について大數觀察なる特殊の研究方法を適用し、その結果をば組織的に整理する以外に途はない。

即ち彼等に從へば、統計學の對象は他の社會科學に於けると同じく常に社會現象である。これは統計學は一箇の實

質的科學だといふ意味である。一般に學問は、その對象に從つて實質的科學と形式的科學とに分類される。例へば物理學又は經濟學はそれぞれ物理現象及び經濟現象といふ實在的內容を對象とするから卽ち實質的科學であり、反之、論理學又は數學は共に思考の方法といふ抽象的內容を對象とするから卽ち形式的科學である。さて社會統計學派に從へばその對象は實在的な社會現象であるから、當然實質的科學の範疇に包含さるべきであるが、この社會現象なるものは社會學・法律學・經濟學・歷史學等の種々の社會科學の對象であるから、これが統計學なる一箇の科學の對象たる爲には、統計學に或る特殊の性質が無ければならぬ。これが卽ち大數觀察法であり、これにより、又これによつてのみ、他の社會科學と區別されるといふのである。

如上の主張に立つ社會統計學派は獨逸を中心とし、從つて獨逸派統計學とも呼ばれてゐる。代表者としてはエンゲル(E. Engel)、エッティンゲン(A. von Oettingen)、ワッポイス(J. E. Wappäus)、ワグナー(A. Wagner)、コンラット(J. Conrad)及び現存のウォルフ(H. Wolff)、ミュラー(J. Müller)、チスカ(C. von Tyszka)、チヂェーク(F. Zizek)を擧げる事が出來よう。併し眞にこの學派の主張たる實質科學として統計學の建設に最大の貢獻を爲したのは、「統計學と社會學」の著者マイル(v. Mayr, 1841—1925)であらう。彼以外の學者は多くは大數觀察法の研究に終始し、寧ろ形式的科學としての統計學に左袒するの結果に終つてゐるのである。

第二の方法學派は統計學を以て實質的科學と解せず、寧ろ統計的研究方法そのものを研究對象とする學問と見做すのである。換言すればこの學派の所謂統計學とは、統計資料の蒐集、整理及び解析の幾多の方法をば組織的に記述したものである。それは一箇の方法學(Methodenlehre)であつて、從つて一箇の形式學(Formalwissenschaft)である。故にこの派に從へば、凡ゆる集團現象に適用されねばならず、決して社會現象にのみ局限される必要はない事になる。例へば生物學・天文學・氣象學その他萬般の領域に於て利用されねばならぬ。大勢はこれに傾いては居るが、併しこの派のうちにも適用の範圍を社會的集團現象に限定する人もある。これについては專門家の間にも最後的結論は發見されてゐない。

方法論學派は右の如く統計的方法卽ち資料の蒐集・整理及び解析の方法を研究對象とするものであるが、實際には殆ど全く最後の解析方法のみに限定されてゐる觀がある。蓋しこの派の理論的基礎は確率理論にあるが、この理論は統計的解析に於て始めてその威力を發揮するからである。斯くて今日の方法論學派は所謂「數學的統計學」によつて代表されてゐると言つて差支へがない。

方法論學派はその代表者が英米に蝟集してゐるところから英米派統計學と呼ばれ、上記の獨逸派統計學に對立してゐる。素より獨逸にも曾てはリューメリン(v. Rümelin)及び特にレキシス(W. Lexis)の如きこの派の先驅者を輩出したが、確率理論の巧妙な適用によつて、この派の基礎を確立したのは、英國の著名な生物學者カール・ピアソン(Karl Pearson)であり、その遺業を繼いだのは同じく英國のエッヂワース(Edgeworth)、ボーレー A. L. Bowley)、ユール(U. Yule)又は米國のムーア(H. Moore)、パーソンズ(W. Persons)等である。猶ほこの學派は他の諸國にも優秀な學徒を獲得し、伊太利のヂニ(Gini)、モルタラ(Mortara)、佛蘭西のアフタリオン(Aftalion)、マルシュ(March)、丁抹のウェステルガード(H. Westergaard)を擧げ得べく、數年前物故せる伯林大學教授ボルトキウィッチ(von Bortkiewicz)は露西亞のチュプロウ(Tschuprow)と共に純理論に不朽の貢獻を爲したのである。我國の

統計學は久しく獨逸派の影響下にあつたが、最近の傾向は英米派に傾き、この方面の進歩は見る可きものがある。併し他の諸科學に比し遜色ある事は否定し得ない。これは我國に統計學的研究に對する眞摯な要求の起つたのが極めて最近の事に屬するからである。

私がこの講義に於て重點を置くのも實にこの統計的方法、就中そのうちの解析法である。調査の問題はその大様を記すに止める。これは調査を輕視するからではなく、第一には紙數の關係からであり、第二には調査は概して吾々の行ふものではなく、吾々は主として既存の資料を利用する立場にあるからである。併し利用の場合に、資料が果して信頼し得るものかどうかは、解析の結果を左右する重要な問題であるから、一應の知識は必要である。羸弱な材料を以てしては、如何に高度の技術を以てしても理想的建造物の得られない事は言ふ迄もなからう。故に資料判斷に必要な程度の、調査に關する知識は除外するわけには行かぬ。

註　統計學史の參考書としては次の二つが勝れてゐる。

H. Walker: Studies in the History of Statistical Method, 1929.

H. Westergaard: Contributions to the History of Statistics, 1932.

第三節　經濟學と統計學との關聯

社會的大量を記述し解析する方法が統計學の獨自の課題である以上、凡ゆる社會科學は、少くとも量的問題を取扱ふ限り、多かれ少かれ統計的方法を援用すべき餘地があり、また必要がある。蓋し如何なる社會科學も、その對象は社會なる集團であつて、決して個々の獨立的事象ではないからである。一見箇別的事象を取扱ふかに見える歴史學も、それが學たる以上は、決して相互に遊離した箇別的事象を羅列するのではなく、實は社會的に規制された人間により、又人間との關係によつて發生した事象を因果的關聯に於て記述するものである。併し社會的事象の總べてが統計的研究の對象たり得るものではない。その對象たり得る爲には、量的性質を具有せねばならぬことは、統計學の本質から明かである。故に統計的研究法の援用され得る程度は、その科學の對象の量的性質が大なるか小なるかに比例する筈である。

この點からすれば、經濟學は凡ゆる社會科學のうち、最も統計的方法に適當した學問だと言へる。何となれば、經濟學の對象たる經濟事象とは何れも價格と關聯せる現象に外ならぬからである。吾々が或る事象を特に「經濟的」なものとして經濟學に採り入れるのは、該事象を價格の觀點から考察するからである。故に同一事象も、もし他の觀點から考察されるならば、既に經濟學の領域には屬さない事になる。例へば「住宅」は建築費・賃貸料又は家屋稅等の見地から觀察されゝばこそ經濟學の對象となるのであるが、もし所有權の見地からすれば法律學の、健康の見地からすれば衞生學の、材料や構造の見地からすれば建築學の對象となる如きこれである。

斯く或る事象を特に經濟學の對象たらしむるものは價格に外ならぬとすれば、その量的表現の容易な事は明かである。蓋し價格はそれ自體が量的なものだからである。凡ゆる社會科學のうちで經濟學が特に統計的方法に適當だといふのはこの意味に於てゞある。然るに實際にこの方法が經濟學に採り入れられたのは比較的最近の事で、前世紀末葉までの永い經濟學の歴史に於ては、殆ど全く關聯を持たなかつたのである。この事は正統學派・歴史學派又は限界效

用學派の研究方法を顧れば容易に肯けると思ふ。正統學派は人間の利己心を抽象化した「經濟人」なる非現實的概念を演繹的に展開する事によつて經濟的眞理へ到達せんと圖つたが、斯かる思辨的方法に於ては數量は始めから問題とならず、從つて統計學は何等の關聯も持ち得ないのである。正統學派に對する反動として興つた獨逸の歴史學派は、經濟事象の具體性・箇別性に囚はるゝ餘り、箇々の事實の記述に終始し、大數現象としての經濟現象を忘却した點に於て、大數現象を問題とする統計學とは縁が遠い。更に限界效用學派に至つては、正統學派の演繹法を更に押し進めたもので、この學派の三つの傾向、即ちジェヴォンスの數學的交換理論、ワルラスの靜態均衡理論及びメンガーの心理的理論は何れも現實の背景を缺く高度の抽象理論であるから、具體的方法たる統計的方法は全く排除されてゐるのである。

然るに統計的研究法の必要を始めて最も明白に唱へたのが、この限界效用學派の偉大な創立者の一人たる前記ジェヴォンスである事は注目に値する。彼はその「經濟學純理」(一八七一年)に於て次の如く言つてゐる。「演繹的な經濟學は純歸納的な統計學によつて證明されまた效果あらしめられねばならぬ。……もし商業統計が今日よりも遙かに完全正確にして、爲に公式が數字的資料の援けによつて正確な意味を與へられるならば、經濟學は次第に精確科學(exact science)の域に近づくであらう。斯かる資料は主として、社會が獲得し消費する財貨の量並びにこれら財貨の交換される價格の正確な勘定から得られる。固より資料を蒐集する爲には費用と勞力を必要とし、且つ人々が報告を提出する事を嫌ふ傾きがあるが、而も是等統計の要請さるべき事には變りがない。測定され記錄される數量それ自身は最も正確精密である。若干の方面に於ては吾人は既に略ゝ完全な報告を持つてゐる。例へば茶・砂糖・コーヒー又は煙草の如き、全部輸入に俟つ商品これである。併し課税されざる、また多かれ少かれ國內に産出される物品については、吾人は消費量に關して未だ極めて漠然たる概念を有するに過ぎない。尤も農業統計の蒐集には幾分の曙光が現はれて來た。そして棉花其他の取引者は在庫量・輸入量及び消費量の正確な報道を得る必要を大いに感じてゐるから、從來吾人の持つよりは遙かに正確な統計がいづれ發表されるに至るであらう」と。彼は理論經濟學の方面に於ては極めて抽象的な數理的理論を展開するに止まつたが、併し物價指數や恐慌に關する研究に於ては、素朴ながらも統計的方法を援用してゐるのである。

經濟學に統計學の適用されなかつた理由は、ジェヴォンスに從へば右の如く資料の缺如にある。併しこれ以外に更に重大な理由として、統計的研究法そのものゝ當時極めて幼稚な階段に在つた事を擧げねばならぬ。既に統計學の歴史に於て說明した通り、統計的解析法に主力を注ぐ方法論派の擡頭したのは主として一八九〇年以後の事である。時系列解析法又は相關理論の如き今日統計學に於て最重要の地位を占める各種理論は、ジェヴォンスの頃には未だ開拓されざる領域であつた。即ち資料の缺如と統計理論及び技術の不備が相俟つて、統計的方法の適用を遲からしめてゐたのである。

今日に於ては事情は異る。一方には資料の蒐集が急速度に行はれ、他方統計理論及び技術の進步は目覺しきものがある。經濟學の研究方法がこれに伴つて急回轉を示すに至つたのは毫も怪しむに足りぬ。私はこゝに詳しくその經過を述べる暇はない。唯だ簡單に次の事だけを指摘して置きたい。即ち統計的方法の效果が一度び認識さるゝと共に、過去の抽象的方法に對する反動として、統計に立脚する現實的方法が急激に擡頭し、その極端なるに至つては「理論」

否定の態度を示すものすら現はれた。一九一九年以降暫く學界に君臨したハーヴァード研究所の統計的景氣觀測の如きその代表的なものである。然るに彼等の意圖は幾何もなく事實によつて否定されざるを得なかつた。事實から出發したものが事實によつて否定されるのは大なる皮肉ではあるが、畢竟統計的方法に對する過信の當然の歸結に過ぎない。即ち統計的事實は資料に規制された經驗に外ならず、常に空間と時間の制限の下にしか妥當し得ないものである。斯かる事實は推論の基礎としては無限の重要さを具有するものであるが、併しそれ以上のものではあり得ない。經濟學が經驗科學であるといふ事は、決して經濟學が常に經驗的事實の領域を出でてはならぬといふ意味ではない。それが一箇の學問である以上、その窮局の目的は現象の基本的形態に到達する事に在る。これは畢竟事實は理論に高められねばならぬ。換言すれば統計的事實は理論の前提だといふ事である。故に經濟學の基本的方法は依然として演繹法たる可き事は明かであらう。唯だこの場合、抽象的非現實的命題より出發するとすれば、最初から經驗科學の本質を沒却する事になるのである。要するに經驗科學たる本質上、經濟學の命題即ち理論は事實より出發すべく、また事實によつて吟味さるべきであり、從つてこゝに經濟學と統計學との不離の關聯を發見し得るのである。

斯かる認識は純統計的方法の失敗によつて最近急激に高められた觀がある。一般にエコノメトリク（Econometric）――計量經濟學――と呼ばれ、又アメリカのムーアによつて綜合經濟學（Synthetic Economics）と命名された輓近の潮流は、要するに純粹の思辨的方法又は純粹の經驗的方法に偏する事なく、兩者を打つて一丸とした渾然たる新體系を求めんとするもので、曾て横行した抽象的な數理經濟學を統計的事實の上に基礎づけんとする試みに外ならない。この潮流に乘つて既に幾多の貴重な研究が發表された。私はこの新體系の將來に最も大なる期待を懸けるものである。

なほ箇々の部門に於て如何に統計的方法が適用さるべきかの具體的問題は、別の「經濟統計學」の課題である。

第四節　研究文獻

統計の理論及び技術に關する著作は最近何れの國に於ても頓に激增した。特殊部門を取扱つた專門的著書乃至論文までを擧げれば限りがないから、次に主として大戰後に出版された英・米・獨・佛及び日本の概論的著書だけを揭げて置く。猶ほ伊太利及び露西亞には極めて優れた著書があるが、こゝには一切省略した。

Secrist- An Introduction to Statistical Methods, 1923
Mills- Statistical Methods, applied to Economics & Business, 1929
Davis & Nelson- Elements of Statistics with Application to Economic Data, 1935
Arkin & Colton- Statistical Methods as applied to Economics, Business, Education, Social and Physical Sciences, 1934
Harper- Elements of Practical Statistics, 1930
Rhodes- Elementary Statistical Methods, 1933
Croxton & Cowden- Practical Business Statistics, 1934
Florence- Statistical Method in Economics, 1929
Winkler- Grundriß der Statistik (I. Band, Theoretische Statistik, 1931)
〃 - Statistik, 1925
Tischer- Grundlegung der Statistik, 1929

Möller- Statistik, 1928
Wolff- Theoretische Statistik, 1926
Müller- Theorie und Technik der Statistik, 1927
Anderson- Einführung in die mathematische Statistik, 1935
March- Principes de la Méthode Statistique, 1930
Darmois- Statistique Mathématique, 1928
〃 Statistique et Application, 1934
Jordan- Statistique Mathématique, 1927
Aftalion- Cours de Statistique
Moncetz- Initiation aux Méthodes de la Statistique, 1935
小倉金之助—統計的研究法（積善館）
田村市郎—經濟統計（日本評論社）
中川友長—統計研究法の基礎（同上）
森 數 樹—統計學概論
宗藤圭三—統計學原理
蜷川虎三—統計學概論（岩波書店）
〃 —統計利用に於ける基本問題（同上）
〃 —統計學研究（同上）
森田優三—統計學概論（森山書店）

小 林 新—經濟統計學（ダイヤモンド社）
汐見三郎—統計學（日本評論社）
松村宗治—數理統計學（内田老鶴圃）
北村友圭—統計數學（高岡書店）
木村敎雄—統計法概論（培風館書店）

同じく統計理論に關する學術雜誌としては次のものを擧げられる。この内には全く統計學專門のものもあるし、また一般社會及び經濟理論に伍して揭げられてゐるものもある。

Journal of the Royal Statistical Society
Journal of the American Statistical Association
Allgemeines Statistisches Archiv
Journal de la Statistique de Paris
Metron
Econometrica
Biometrika
Revue de l'Institut International de Statistique
Jahrbücher für Nationalökonomie und Statistik
Zeitschrift für Nationalökonomie
Journal of Political Economy

Quarterly Journal of Economics
Archiv für Mathematische Wirtschafts-und Sozialforschung
Annals of Mathematical Statistics
統計集誌
統計學雜誌
日本統計學會年報

第二章 統計調査

第一節 統計調査の諸問題

統計學の對象が個體に非ずして集團なる事は上に述べた。こゝに言ふ集團とは或る標識によつて客觀的に規定された等質性個體の集合であつて、これを統計集團と名づける。この統計集團を數字的に表現したものが統計であり、統計に關する組織的知識が統計學なのである。即ち統計學の出發點は統計集團である。集團は素より個體の集合であるが、この集合は主觀的に行はれる事もあり、又客觀的に行はれる事もある。統計集團とは後者のみを指す。そしてこの場合には、集團を構成する各個體（統計單位）は何等かの共通性即ち等質性を持たねばならぬ。全く異種の個體は、如何なる手段を講じても、これを相互に結合する事は不可能だからである。人口統計の對象たる人口集團は生物學的に規定された人類の範疇に屬する個人即ち單位の結合されたものであり、また婚姻統計の對象たる婚姻集團は、法律的に規定された夫婦關係に屬する個體即ち夫婦の結合されたものである。馬と牛とは全く別箇のものであるから、馬匹統計の對象たる馬匹集團から牛は除外される。併し家畜統計となれば、馬も牛も共に家畜たる共通點を持つから、同一集團に結合されるのである。

斯く統計集團は同種の個體の結合されたものであるが、同時にこれら各個體は相互に或る相違點を持つ事を必要と

する。人口集團を構成する各個人が、もし總ての點に於て全く等しかつたとしたら、何を苦しんで統計的研究を施す必要があらう。然るに實際には各人は人種・國籍・年齢・職業等幾多の點に於て相互に相違してゐる。斯かる相違點こそ統計的研究の對象となるのであつて、これを調査事項或ひは標識といふ。

即ち吾々が統計集團を把握するには、該集團を構成する各單位につき、一定の標識を數へ上げねばならぬ。この手續を統計調査といふ。統計調査には國勢調査の如き大規模のものから、少數の個人で行ふ小規模のものに至るまで、大小様々の階段があるし、又本來の意味の統計調査の外に各種の代用的な調査法もある。調査に於ては技術的問題が甚だ多いが、こゝには單に原則的な若干問題を記述するに止める。

註　詳しくは專門書について見られたい。戸田貞三「社會調査」又は Franz Zizek : Wie statistische Zahlen entstehen, 1937 なぞは手頃な手引きである。

（一）單　位

等質的個體の集合が統計集團を構成するならば、斯かる個體は統計的單位である。即ち吾々は統計的單位を結合する事によつて統計集團を決定し得るのであつて、從つて統計調査の第一前提は單位の確定になければならぬ。單位を大別して計量單位と計數單位に分ち得る。物の大さ・重さ・價値・年齢等の「量」は度量衡、貨幣又は年月の如き客觀的尺度によつて測られるから、單位の決定に格別の問題は起り得ない。唯だ同一の名稱を持ち乍ら異る內容を有する尺度があるから、それを明かにして置く必要はある。例へば標準に尺を用ひたとき、曲尺なるか鯨尺なるか、噸を用ひたとき米噸なるか英噸なるか、或ひは年齢を示す場合に數へ年なるか滿歳なるか等を明瞭に規定して置かねばならぬ。

これに對して事物の箇數を數へ上げる場合の計數單位については多くの問題がある。單位が自然的に與へられてゐる場合（例へば人口調査に於ける「人」、馬匹調査に於ける「馬」等）には問題は簡單である。然るに社會的人工的存在物又は人の行爲等の調査に於ては、單位は著しく曖昧となる。これは對象物の一律の定義が困難だからであつて、例へば工場・住宅・學生・婚姻・失業等の調査に於ては、先づ工場とは何か、失業とは何か等を豫め嚴格に定義しない限り、單位の決定される筈はない。然るに例へば失業について見るに、一般に失業とは「勞働の意思と能力を持ち乍ら勞働の機會を得ざること」と定義されるが、果してこの定義で滿足な調査が行へるであらうか。この定義に從へば、未就職者の大なる部分は失業者であり、反之、所謂「半失業」即ち臨時的にしか雇傭されない勞働者は失業者ではない事になる。要するに單位の問題は定義の問題から解決されねばならない。故に卒然二箇の失業統計を比較せんと欲しても、無意味に終る場合が多い。統計は信用し得ないと屢々言はれるが、その主たる理由の一つは實にこゝに在るのである。

（二）範　圍

人口調査に於て地球上の凡ゆる人を調査するとか、物價調査に於て凡ゆる商品價格を調査するが如きは言ふべくして行ひ得るものではない。故に一國の人口とか重要商品の價格とかに範圍を限定するのである。併し調査範圍の限定は單に調査の可能不可能の問題から生ずるばかりでなく、また調査の目的からも當然必要となつて來る。吾々の行ふ凡ゆる研究に於て、研究範圍は豫め限定されて居る筈であつて、その範圍に何等の限定もない研究の如きはあり得な

い。統計調査は何等かの研究目的の爲に行はれるのであるから、研究目的は必然に統計調査の範圍を限定する事明かであらう。

即ち統計調査は研究目的に必要な統計集團に就て行ふ可きである。そして大數觀察の原則から、この統計集團の各單位全部につき、所定の標識を調査するのが當然で、これを悉皆調査と言ふ。併し各單位全部の代りにその一部について調査し、所謂部分から全體を推すの方法を採る場合が尠くない。これを代用調査といふ。その種類・方法は次節に述べる。

範圍の限定は場所的・屬性的及び時間的の三種類に分れる。場所的限定とは一定の地理的區域に限つて調査する事で、國勢調査は全國につき、失業調査は大都市及びその附近につき、また日銀物價調は東京市だけについて行はれるが如き、この例である。國勢調査が全國について施行されるのは、目的上當然であるが、失業調査が都會地に限られるのは、我國でいふ失業は工業勞働者又は自由業者の如き主として都會的職業の失業を指すが爲である。物價調が同じく都會地に限定されるのは、元來物價は都會地の商品市場又は卸店に於て決定されたものが全國の基準となる性質のものだからである。次の屬性的限定とは物價調に於て重要商品のみの價格を調査するとか、商工省工場調に於て「常時五人以上の職工を使用する工場」のみを調査するとかの類である。右の意味の工場と工場法に於ける工場（職工十人以上）とは從つて別の概念なる事に留意せねばならぬ。最後に時間的限定は或る一定時點に於ける調査と、一定期間内の繼續的調査とを區別せねばならぬ。國勢調査は十月一日午前零時の状態を對象とし、反之、出生・死亡・婚姻とか輸出入額とかの調査は一箇月、三箇月、六箇月、一箇年等の期間に亙つて行はれる。前者を靜態統計、後者を動態統計といふ。統計の扱ふ量は時間の經過と共に變化する所謂「動大量」であるが、一瞬時を取て觀察すれば恰も不動的確定的大量即ち靜大量と認められる。人口は出生・死亡によつて絶えず變化してゐるが、某月某日某時の人口は何萬何千人と確定される。大量の構成を分析するには斯かる靜大量について行はざるを得ない。蓋し上述の如く大量は不斷に增減し、從て構成狀態も絶えず變化してゐるからである。斯くて靜態統計の目的は實にこの構成狀態の分析に在るのである。國勢調査に依て始めて人口の年齡別・職業別・體性別等が判明する。反之、大量の變化を知るを目的とするものが動態統計で、一定期間繼續して觀察が行はれる事を必要條件とする。固より靜態統計と雖も、單に一回の調査で終る事は殆どなく、一定の期間を置いて繰返へされるのが普通である。國勢調査は五年毎に、物價調は毎月、日銀兌換券發行額は毎日繰返へされる。これを相互に比較すれば、時の經過に伴ふ構成の變化も知り得るわけであるが、而もこれは斷續的なもので、調査の行はれなかつた中間期の計數は補間法によつて推察し得るに止まる。この點、連續的な動態統計とは全く性質を異にするのである。

（三）標　識（調査事項）

單位と範圍とが決定されたならば、次には調査す可き事項を決定せねばならぬ。既述の如く、單位は等質性を有するが故に統計集團に結合されるのであるが、併しその意味は、各單位は相互に共通性を有し、これが爲にこれら各單位が一箇の範疇に集約されると言ふだけで、決して各單位が總ての點で全く等しいと言ふのではない。否、各單位が相互に共通點の外に更に相違點を持つといふ事が統計調査を可能ならしめると同時に必要ならしめるのである。蓋し全く共通性が無ければ統計集團は成立しないから、當然統計調査は不可能となり、又もし全部が總ての點に於て等し

いならば統計調査の必要はなくなる筈である。國有鐵道の線路の幅とか煙草の小賣價格とかを特に統計的に研究する人はあるまい。即ち單位の持つ相違性こそ統計調査の對象となるのであつて、これを調査事項又は標識といふ。例へば人口なる統計集團を構成する各單位即ち各個人は國籍・人種・年齢・體性・職業等の點に於て相互に相違してゐるから、これらは皆標識である。かゝる標識の或るもの(一つ又は數箇)を選擇し、これを各單位について調査するのが統計調査なのである。

標識も亦屬性的・場所的及び時間的の三種に分けられる。最も決定に困難を感ずるのは第一の屬性的標識であつて、體性や身長の如く客觀的決定の容易なものは別として、職業の如き屬性は限界が屢ゝ不明瞭で、爲に誤りに陷る場合が尠くない。

一回の調査に際して幾箇の標識を調査すべきかは勿論調査目的によつて異る。併し大規模の調査ほど標識の數を減少する必要ある事は忘れてならない。蓋し餘りに複雜な質問に對しては答に誤りも多くなり、時には煩を嫌つて調査に應じない人も出て來るからである。强制的な國勢調査に於てすら、標識の極めて少い事實は、自發的好意に俟つ一般の統計調査に於て、大いに參考とされなければならぬ。

最後に注意すべきは、「單位と標識」とは統計集團の採り方如何によつて隨時に決定されるもので、決して先天的に與へられてゐるものでないといふ事である。例へば「工場」は工場調査では單位であるが、建築調査では標識に過ぎず、「自殺」は死亡調査では標識であるが、自殺調査では單位となるが如き、この類である。これは要するに如何なる集團的事物も、同種の集團と認める事も出來るし、又より大なる集團の一部とも認める事が出來るからである。換言すれば、吾々の構成する「概念」の問題に歸着する。吾々は個別現象間の類似に基いて概念を構成するのであるが、概念を確定せんとすればするほど、包含される範圍は狹くなるものである。人間、日本人、日本男子、軍人、陸軍々人、大佐といふやうに、廣い概念はそれが確定されるに從つて狹い概念となつてくる。これを更に押進めれば最後には、陸軍大佐某氏といふ個人に歸着して了ふ。この間の樣々の概念によつて樣々の集團が構成されるのであるから、その都度從來の標識は單位に變つて來る。逆に見れば單位が次第に標識に變つてゆく。先驗的な單位とか不變的な標識とかの有り得ない事は、これによつて明かであらう。

(四)調査の障碍

統計資料が信憑し得るか否かは、統計的研究の成果を左右する根本問題である。然るに調査には各種の障碍が起り易く、從つて完全無缺の資料なるものは先づ得られないと言つてよいのである。故に資料利用者は、起る事あるべき障碍の性質や程度を考慮して、以て研究結果に多少の斟酌を加へねばならぬ。

調査の障碍は調査者の側より生ずるものと、被調査者の側より生ずるものとがある。調査は多數調査者の協力によつて行はれるのが普通であるが、この場合彼等の間に調査の單位・標識・範圍又は調査方法等に關する誤りが起らぬとは限らない。毎回の國勢調査に際して豫め調査員に調査心得を配布したり、豫習を施行したりするのは、要するに是等についての誤謬を避ける爲である。併し一般に調査者は自身が專門家であるか、乃至は專門家の監督下に在るから、彼等の側より生ずる誤謬は少いと見てよい。

反之、被調査者の側より生ずる誤謬は時に極めて甚大に達し、調査をして全く無效に終らしめる場合も少くはない。

この場合の誤謬は無意識的なものと意識的なものとがある。被調査者が調査事項を充分に了解しない事から、誤つた報告を爲す事例は最も多からう。併し調査の主たる障碍は寧ろ意識的誤謬に在るのである。蓋し調査の種類によつては被調査者の利害や個人的秘密に立ち入る場合があるが、これについて率直な報告を求め得るか否かは疑問であらう。財産又は所得の調査は、被調査者が課稅と結びつけて考へるため、概して內輪に申告される恐れがある。故に稅務署では生活狀態から推して或る程度の修正を施す必要もあるわけである。前科の有無、家庭の內情、若い婦女子の年齡等は何れも個人の名譽・祕密又は虛榮心に關する事項であるから、動もすれば虛僞の報告を受け易い。我國の國勢調査で年齡欄に生年月日を記入せしめるのは、虛僞の申告を防止する一方法である。

要するに調査の障碍のうちには調査者の努力親切によつて排除されるものがあると共に、問題の性質上、如何とも爲し難いものもある。斯かる事項の調査は始めから結果の不正確を覺悟してかゝらねばならず、利用者も亦特に愼重を必要とする次第である。

(五) 調査の方法

統計調査は調査の對象たる統計集團の各單位について、直接且つ總括的に行ふ可きものであるが、斯かる操作は、對象の範圍が廣大な場合には時間・勞力・費用の點からして常に必ずしも行へるものではない。これに代るものが間接調査と代用調査である。

間接調査とは既に他の目的の爲に蒐集された資料から調査事項を求める方法を言ふ。人口統計について見るに、定期的に行はれる國勢調査は人口の數及び構成に關する統計を求める目的で全人口について直接に行ふ調査であるが、反之、出生・死亡・婚姻等の統計は何れも屆出の結果を利用した間接的調査に過ぎない。斯く調査が直接的なるか間接的なるかによつて第一次統計（Primary statistics）と第二次統計（Secondary statistics）に區別される。吾人の利用する大部分の統計は寧ろこの第二次統計と言つてもよいのである。蓋し吾人は直接に調査對象とされる事は稀なるに反し、間接的には不斷に資料の供給を行つてゐるからである。例へば出生や結婚を屆出でれば人口動態統計に、物品を生產・販賣又は消費すれば各種の經濟統計に、法を破り又は自殺したりすれば道德統計に、それぞれ記錄を留めるわけで、これらは何れも吾人の識らぬ間に何時しか統計に纏め上げられて了ふのである。

斯かる間接調査に於ては、調査洩れの危險の多い事は勿論であるが、それ以外に一つの大なる缺陷がある。それは例へば婚姻の調査を婚姻屆出から作製する場合には、單に法律的婚姻しか求められない。事實上の夫婦でも、屆出をせぬ限りは、法律上の夫婦ではなく、從つて婚姻統計には現はれて來ない。國勢調査では「配偶關係の記入は實際の狀態に依るもので、必ずしも戶籍と同一でなくてもよい」とし、間接調査の缺陷を補つてゐる。

次に代用調査とは一部の調査から全體を推す方法を言ふ。推計、部分調査、標本調査、アンケート等がその主なものである。（一）推計とは例へば我國の人口は各五年每に行はれる國勢調査によつて判明するが、その間の各年の人口は斯かる國勢調査の結果から割り出す外はない。今昭和八年の人口を求めんとすれば、國勢調査の行はれた昭和五年及び十年の兩人口から右五年間の年平均增加率 x を算出し、複利算によつて　昭和八年人口＝昭和五年人口×$(1+x)^3$ とするのである。もし增加率が實際に每年等しかつたとすれば、推計は實際と合致する筈である。換言すれば推計が正確な爲には、基礎數とそれに乘ぜられる數との間に正確な量的關係がなければならぬ。もし每年の增加率に大なる

相違があれば、右の推定は實際とは懸け離れたものになる。廣く行はれる推計としては、米の收穫高を求める爲に、一定面積の收穫高（坪刈）に耕地面積を乘ずる場合が擧げられる。併し收穫は場所によつて大差があるから、結果は可成り杜撰たるを免れない。

（二）の部分調査とは、一部は全體の縮圖であると想定し得る場合に、その一部を調査する事によつて全體を推す方法を言ふ。例へば國勢調査の結果を整理するには相當の長期間を要するが、抽出調査なる部分調査によつて大體の結果が速報される。併し部分が全體の縮圖と認められる爲には、可成り大きな部分でなければならぬ事は言ふ迄もない。（三）の標本調査とは同じく部分調査ではあるが、異る點は、第一に豫め典型的な單位を選定すること、及び第二にその選定した單位について充分詳しい調査を行ふ事である。勞働者家計調査はこの方法に據つて生計費の内容を明かにせんとするものであるが、調査される家族が果して典型的かどうかは調査者の主觀によつて決定されるほかはないから、結果の妥當性に就ては多分に疑問の餘地がある事になる。（四）のアンケート（Enquête）とは調査事項に特に關係を持つ人々に直接口頭又は書面なぞで問ひ合はせる方法を言ふ。意見の聽取が主たる目的であつて、必ずしも數字的解答に限られないから、統計調査とは言へないかも知れない。アンケートとは元來「疑はしい事項を闡明する爲に證言を集める」といふ佛蘭西語である。

調査事項記入の方法は、被調査者が自ら記入する自計式と、調査員が記入する他計式とに分たれる。調査事項が簡單ならば自計式が便利で、我國の國勢調査はこの方法に依る。併し調査事項が複雜な場合、又は被調査者の知識程度が低い爲に正確な申告を期待し得ない場合には、他計式によらざるを得ない。

第二節　統計機關と統計資料

統計機關　統計調査又は編纂を主たる任務とする機關、又は統計理論の研究に從事しつゝある機關を統計機關といふ。官廳的、私人的、或ひは國際的等に分ける事が出來る。先づ官廳的機關について見るに、一國の國土・資源・人口・産業等に關する正確な統計が政治の不可缺的前提たる事からして、政治團體にして自ら統計機關を所有しないものは考へられない。小は市町村から大は國家に至るまで必ず何等の形態の統計機關を具へてゐるのである。いま日本の中央政府に就て見るに、各省に獨自の統計課あるは勿論、更に內閣直屬の內閣統計局なる厖大な統計機關がある。その掌る事務は、（一）行政各部統計の統一、（二）國際統計事務の統轄、（三）人口統計・勞働統計其他國勢の基本に關する統計にして行政各部に專屬せざる事項、（四）統計に關する圖書の刊行、其他である。そのうち最も主たる任務は言ふ迄もなく國勢調査であつて、一定日時を期して各國民を一齊に調査するが如き大規模の事業は、國家的資力と權力を背景にせねば不可能であらう。

次の國際的統計機關には官廳的なものと私人的なものとがある。國際關係が密接となるに從つて、國際的統計の作製は益ゝ重要を加へ來つた。官廳的なものとしては萬國電信同盟及び萬國郵便同盟、國際農業局又は國際聯盟の如き、國家として加入してゐる團體に附屬せる統計機關がある。うち最も重要なのは、國際聯盟に屬する國際勞働事務局であらう。國際的私人機關としては國際統計協會がある。曾てケトレーによつて創設された國際統計會議（Congrès International des Statistiques）は最初の國際的施設として多大の功績を殘したが、一八七八年に解散され、一八八

五年にこれに代つて現はれたのが右の國際統計協會（Institute International de Statistique）である。各國の統計學者及び實際家が相互に意見を交換し、國際統計の進歩を計るを目的とし、既に二十數回の大會を開いてゐる。昭和五年には東京に開かれ、全世界の著名な統計學者が參集した事は記憶されてゐる方もあらう。

最後に國内的の私人研究團體としては英國の王立統計協會、米國のアメリカ統計協會、獨逸の獨逸統計協會、佛蘭西のパリー統計協會等が最も有名である。何れも優秀な機關雜誌を發行し、統計理論の發達に貢獻するところ大である。我國には日本統計學會、帝國統計協會の兩者が代表的なものと言へよう。これ以外に、各種新聞社、經濟雜誌社、會社及び銀行、株式取引所、又は商工會議所等に屢〻完備した統計機關がある。これらに就ては次の「統計資料」の章を參照されたい。

統計資料　各種統計機關は不斷に統計調査乃至整理に沒頭し無數の資料を提供しつゝある。作製者の如何に從つて官廳統計と私人統計とに分ち得べく、對象の如何に從つて人口統計・經濟統計・道德統計・醫學統計・氣象統計等々に分ち得る。こゝには本邦に於て廣く利用される經濟資料の一端を擧げて置く。

（一）年刊資料

日本帝國統計年鑑（内閣統計局）
日本帝國統計摘要（同）
人口動態統計（同）
人口動態記述編（同）
日本帝國人口動態統計摘要（同）
家計調査報告（同）
外國國勢要覧（同）
勞働統計要覧（同）
金融事項參考書（大藏省）
外國貿易年表（同）
農林省統計表（農林省）
米穀統計年報（同）
商工省統計表（商工省）
工場統計表（同）
會社統計表（同）
職業紹介年報（内務省）
手形交換所年報（東京手形交換所）
本邦經濟統計（日本銀行）
全國公債社債明細表（日本興業銀行）
日本勞働年鑑（大原社會問題研究所）

國際年鑑（國際文化協會）
ダイヤモンド經濟統計年鑑（ダイヤモンド社）
東洋經濟經濟年鑑（東洋經濟新報社）

（二）月刊資料

賃銀月報（内閣統計局）
勞働時報（内務省社會局）
勞働統計（日本銀行）
東京物價調（同）
東京小賣相場調（同）
銀行會社計畫資本調（同）
銀行通信錄（東京銀行集會所）
東京手形交換所月報（東京手形交換所）
重要株式利廻調（日本勸業銀行）
重要債券利廻調（同）
拂込金調（同）
銀行會社計設擴張資本調（三井銀行）

全國倉庫貨物現在高表（三菱倉庫會社）
重要經濟統計月報（神戸商大、商業研究所）

（三）臨時刊行資料

國勢調査報告（内閣統計局）
勞働統計實地調査報告（同）
大正十四年失業統計調査報告（同）
大正十五年乃至昭和二年家計調査（同）
昭和四年農業調査結果報告（同）
大正十四年に於ける國民所得（同）
第五回生命表（同）
明治大正國勢總覽（東洋經濟新報社）
明治大正財政總覽（同）
日本經濟統計總觀（朝日新聞社）
帝國資源總覽（資源局）
商品年鑑（國勢社）

右の外に各種の經濟雜誌、年鑑等何れも多くの統計資料を揭げてゐる。主たるものは、東洋經濟新報・ダイヤモン

ド・エコノミスト・國勢グラフ・朝日年鑑・日々年鑑・時事年鑑・日本經濟年報等である。各種新聞の經濟欄は最も新らしい資料を必要とする人々には缺く可からざるものである。猶ほ外國の資料に就てはこゝでは一切省略せざるを得ない。唯だ英國に Statistical Abstract of the United Kingdom, 米國に Statistical abstract of the United States, 獨逸に Statistisches Jahrbuch für das Deutsche Reich なる總括的統計年鑑あるを記して置く。資料の具體的問題は、別の「經濟統計學」に於て詳述するつもりである。

註 官廳統計の種類は「官廳刊行圖書目錄」に網羅されてゐる。

第三章 整理と圖示

第一節 分類

統計單位を一定の標識に從つて調査した成果は、厖大雜然たる原票の堆積となつて現はれる。標識は單位の相違性であるから、各々の原票は單に單位間の差異を示すに止まる。故に次の手段として、斯かる差異を或る標準に從つて比較的少數の群卽ち階級に分類せねばならぬ。蓋し之によつて始めて差異が類似に要約され、斯くて原票の堆積の內に潛む眞の意味が明瞭となるからである。分類によつて整理されたものを統計系列といひ、統計的研究の基礎をなすのである。

分類の種類とそれによつて得られる統計系列の種類を表示すれば次の如くである。

（一）時間的分類……………時間的系列（時系列）
（二）場所的分類……………場所的系列
（三）屬性的分類 { 量的分類 / 質的分類 }……屬性的系列 { 量的（度數分布） / 質的 }

度數系列
屬性系列

時の經過と共に如何に數字が變化するかを見る爲には、先づ時間的順序に分類せねばならぬ。時間的順序が日、週、

月、年等によつて示される事は言ふ迄もない。次の場所的分類は數字の地理的分布を見る爲で、一般に國別、府縣別、市町村別の如き行政的區劃に從ふ。併し選擧統計に於て選擧區を分類標準とするが如き特殊の場合もある。最後の屬性的分類とは統計單位の持つ性質に施す分類であつて、この場合例へば人々の男女別の如く、體性なる質的屬性によ

(1) 時間的分類

本邦麥酒釀出果年

		石
昭和	2	37,303
	3	41,017
	4	39,156
	5	38,634
	6	36,637
	7	68,812
	8	13,373
	9	118,009
	10	135,107
	11	132,503

(2) 場所的分類

全國工場數府縣別　(昭和10年)

北海道	2,468	滋賀	666
青森	288	京都	3,344
岩手	355	大阪	12,656
宮城	798	兵庫	4,279
秋田	338	奈良	695
山形	867	和歌山	1,380
福島	386	鳥取	332
茨城	866	島根	561
栃木	846	岡山	1,579
群馬	1,786	廣島	2,056
埼玉	1 669	山口	753
千葉	599	德島	417
東京	13,121	香川	510
神奈川	1,382	愛媛	1,229
新潟	1,995	高知	610
富山	640	福岡	1,596
石川	1,645	佐賀	419
福井	2,557	長崎	603
山梨	677	熊本	542
長野	1,438	大分	434
岐阜	1,392	宮崎	322
靜岡	2,983	鹿兒島	903
愛知	9,138	沖繩	74
三重	1,028	合計	85,222

(3) 屬性的分類

(A) 量　的

職工別紡織工場數
(昭和十年末現在)

職工數	工場數
5— 9	11,836
10— 14	3,555
15— 29	4,782
30— 49	2,173
50— 99	1,565
100—199	821
200—499	484
500—999	225
1000 以上	121
計	25,562

(3) 屬性的分類

(B) 質　的

本邦職業別人口 (單位千人)
(1930年)

總人口	64,450
有業者	29,620
農業	14,140
水產業	547
鑛業	251
工業	5,700
商業	4,478
交通業	1,108
公務自由業	2,044
家事使用人	781
其他有業者	571
無業者	34,830

つたときと、又人口の年齢別の如く、年齢なる量的屬性によつたときとが區別される。右の實例に就て知られたい。

これらの分類より生ずる統計系列を見るに、場所的系列及び屬性的系列中の質的系列に於ては變數が一箇なるに對し、時系列及び屬性的系列中の量的系列（これを度數分布といふ）に於ては變數は二箇である。變數とは研究範圍内に於て各種の値をとる數をいふ。場所的に分類された北海道とか青森縣とかの各地方、又は屬性的分類のうちの質的なもの、例へば前例(4)の各種職業なぞは元來數量ではないから變數でない事は言ふ迄もない。反之、前例(1)の昭和二年・三年等は一年たつ毎に殖えてゆく數であるから變數であり、同樣に(3)の(A)に於ける一工場内に於ける職工數による分類は、第一の階級では五人乃至九人、第二の階級では十人乃至十四人といふ風に階級毎に數値が變つて來るから、これ亦變數である。即ち時系列と度數分布に於ては、分類標準とそれに對應する數字が共に變數であり、換言すればこの二つの系列は共に二變數間の數量的關係を示してゐるのである。斯かる關係を數學では函數關係といふ。時系列と度數分布が統計的研究に於て特に重きを爲すのは、函數關係なるものが高度の數學的操作を可能ならしめるからに外ならない。次に述べる圖示法に於ても變數一箇の系列は極めて素朴な圖表でしか示し得ない。統計解析法に於て私が取扱はふとするのは、この變數二箇又はそれ以上の時系列及び度數分布である。

第二節　比　率

數字は常に物の大小の尺度であり、從つて相互に比較される事によつて始めて意味を持つのである。單獨孤立の數字は全く無意味である。甲の身長が五尺六寸だとしても、それが何を意味するかは、例へば他の或る人の又は一般の

身長と比較されて始めて判明する。統計は數字的記述であるから、右と同様に常に比較を必要とする事は言ふ迄もない。統計系列の作製に際しても、絶對數の代りに比較數を用ふる場合が尠くない。併し比較には差を見る場合と、割合（比率）を見る場合とがある。前者は同一種類の數字についてしか行へないに對し、後者は同一種類のみか、更に異種の數字についても行ひ得る。東京府の人口（六百三十七萬人）と大阪府の人口（四百三十萬人）が與へられてゐれば、前者から後者を引く事によつて、東京府の人口が二百〇七萬多い事が判る。反之、東京府の人口とその面積（二、一四五平方粁）が與へられてゐる場合に、この人口と面積を比較するには、面積に對する人口の割合（人口密度）を見る外はない。即ち面積を分母、人口を分子とする割算によつて、一方粁當り二千九百七十人なる答を得るのである。

比較に於て差が重要か比率が重要かは固より斷言し得ない。兩者は同じ程度に重要だと言ふのが最も正しからう。併し差は算出が容易なのに對し、比率はそれが困難であるから、統計作製者がこの困難を引受けて利用者の便を計る事が極めて望ましい。

慶應義塾大學
本科學生數

學部名	學生數	%
文學部	200	6.4
經濟學部	1,661	53.3
法學部	798	25.7
醫學部	455	14.6
合計	3,114	100.0

比率はその構成上三種に分たれる。構成比率・指數及び混成比率これである。構成比率とは與へられた合計が、これを構成する各階級間に如何なる割合で分配されてゐるかを示すもので、普通は合計を一〇〇とし、各階級をそれに對する%として表はす。即ち上例に於て、例へば文學部學生數は $\frac{200}{3114}\times 100=6.4\%$ となる。斯かる構成比率を圖示する最も適當な方法としてパイ圖表なるものがある（次章參照）。

次に指數とは二つ又はそれ以上の階級の大さが與へられてゐるとき、そのうちの一つの階級を基準とし、他の階級をこの基準に對する比率として示したものである。これを廣義の指數といふ。前例に於て文學部學生數を基準とすれば、經濟學部は $\frac{1661}{200}=8.305$、法學部は 3.99、醫學部は 2.275 となり、基準に對して他の階級が何倍に當るかを明示する。併し指數が最も廣く利用されるのは、物價指數・生計費指數の如く、時間的變化を比較する場合である。即ち時系列に於て或る時點を基準とし（基準時點）、他の時點に於ける大さを、基準時點に於ける大さに對する百分比として示すものである。この狹義の指數に就ては幾多の問題があるから、私は別の章で詳論する。猶ほかゝる狹義の指數に對し、廣義の指數を對級比率と呼ぶ人もある。

本邦麥酒釀造能力及釀造高

（單位　千石）

年度	釀造能力	實釀造高	操業率
昭和 1	1,510	814	5.35
2	1,580	803	5.11
3	1,700	894	5.26
4	1,700	905	5.35
5	1,700	821	4.83
6	1,700	763	4.49
7	1,750	767	4.38
8	1,800	1,005	5.58
9	1,800	970	5.39
10	1,840	1,049	5.70
11	1,840	1,211	6.58

最後に混成比率とは異種の數字の比較であつて、前述の人口密度はこの例である。人口と死亡數を比較した死亡率、人口と婚姻數を比較した婚姻率の如きもこれに屬する。上表の操業率なるものも、生産能力と生産高の比率に外ならぬから、矢張り混成比率である。猶ほ混成比率に於て、分母の中に分子に無關係の要素の介在してゐる時には粗製比率といひ、これを除去したときには精製比率といふ。例へば普通の出生率とは、人口千人につき一年間の出生數を指すものであるが、この千人の人口の中には出産に無關係の小兒・老人・男子等が

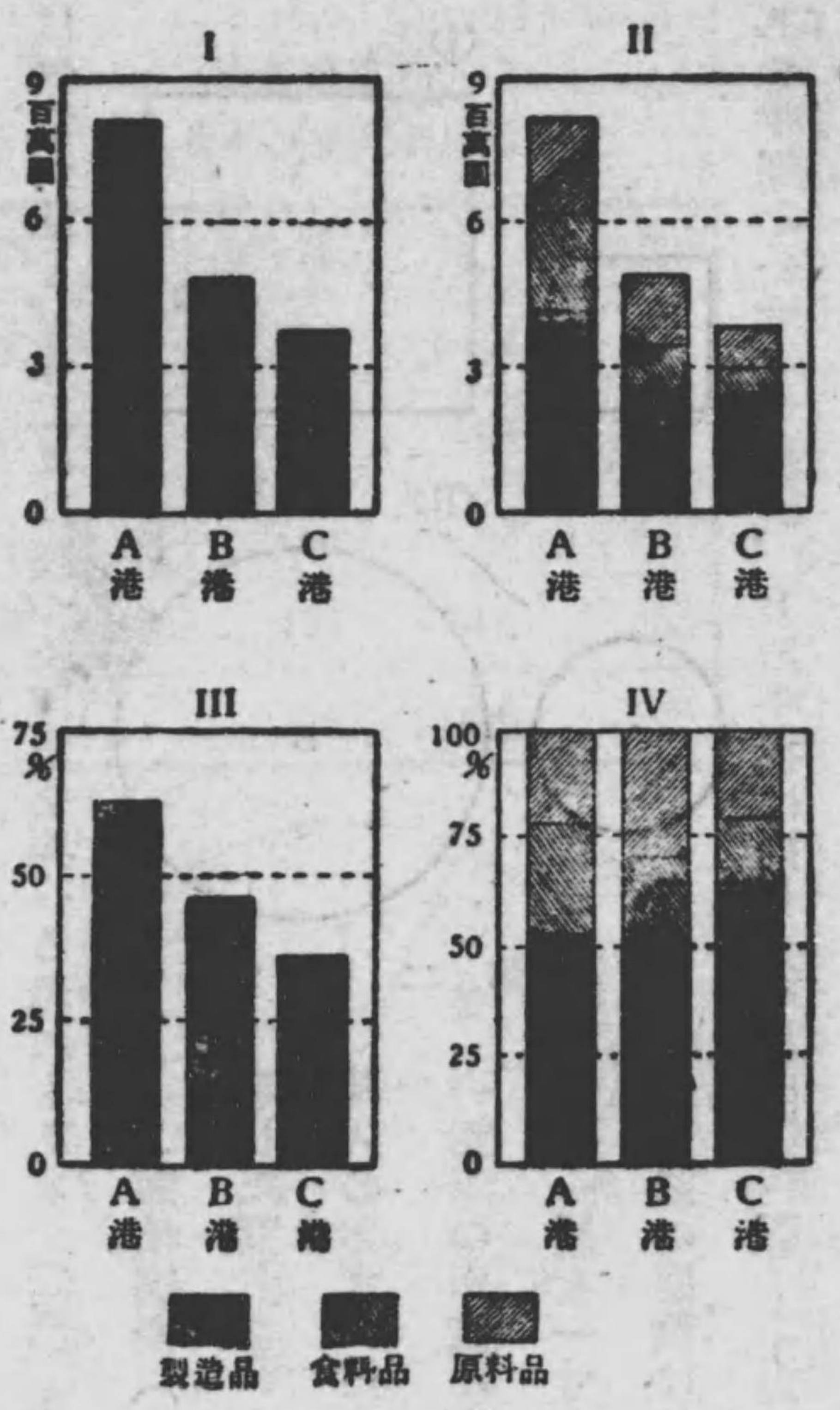
含まれてゐるから、特に姙孕可能年齡の婦人千人に就ての出生數を求める場合がある。前者は出生粗率、後者は出生精率である。人口の年齡別又は性別構成の異る二國の出生率の比較とか又は一國の將來人口の豫測とかに於て、如上の精率は缺く可からざる標準とされる。併しこれを求めるには構成を示す正確な資料を豫め必要とするから、人口統計の整備しない國に就ては算出は固より不可能である。國勢調査の必要はこの點からも充分に認められよう。

第三節　圖示法

統計調査から生ずる亂雜な原票の堆積は、分類によつて整然たる統計系列に配列される。吾人は斯かる統計系列を一覽する事によつて、統計調査の結果をば簡單に讀み取る事が出來るのである。併し數字を讀むこと、即ち數字を比較するには、頭の中で一種の計算を必要とするわけで、動もすれば人々に歡迎されない恨みがある。故に單に統計資料を提供するだけが目的ならば、統計系列を數字のまゝ掲げて置いて足りるであらうが（各種の統計年鑑や內閣統計局刊行物等これに屬する）、統計的事實を强く人々に印象づけるのが目的である場合には、右の如き發表形態は必ずしも適當とは言へないのである。各種の統計圖表が案出され利用されてゐるのは實にこの用意に出づるもので、展覽會の出品物、廣告宣傳のポスター乃至著書論文の中に極めて頻々と現はれるのは之が爲である。統計發表の主要な形態として、統計圖表に關する一應の理解は極めて必要である。

統計系列の分類を說明するに當つて、變數の數如何が分類の一標準であり、變數が一つの場合は場所的系列と質的屬性系列、變數が二つの場合は時系列と度數分布である事を明かにした。圖表化の場合に、この分類は矢張り圖表の種類を決定する標準となるのであつて、變數が一つのときは棒圖、面積圖、立體圖、繪畫圖、統計地圖の如き形式をとり、變數が二つのときは綫圖の形式をとるのである。一般に圖示法は多分に技術的要素を含み、且つ作製者の創意に出づるものが少くないから、こゝではその主たるもののみに就て簡單に說明するに止める。

(A) 變數が一つの場合



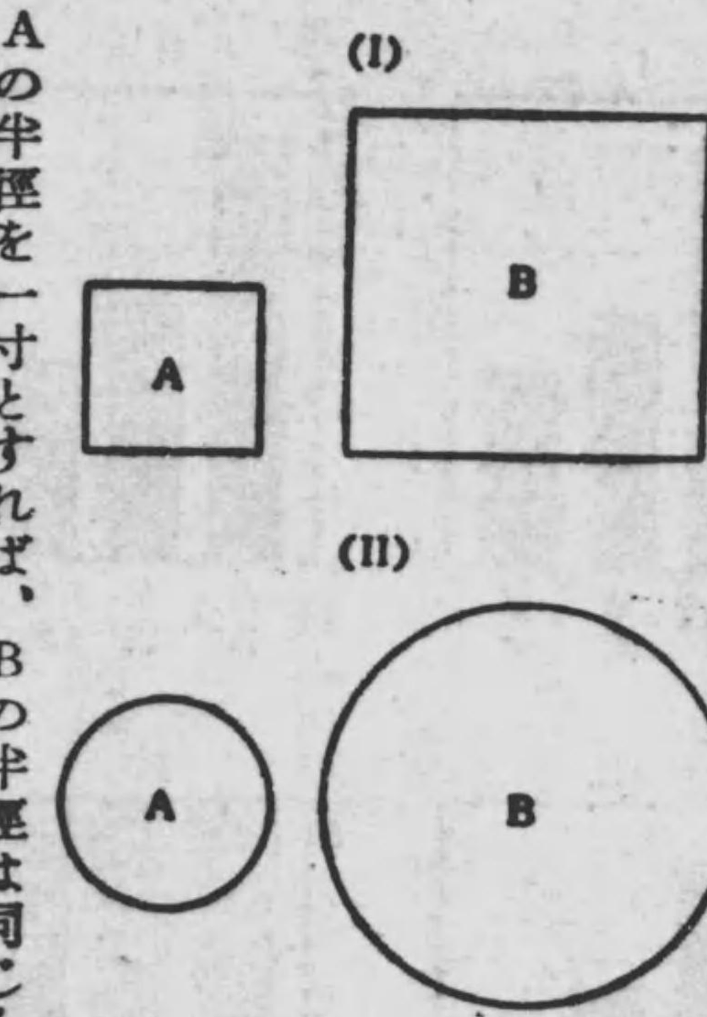
第一圖

(1) 棒圖表（Bar Chart）　數字を棒の長さにて示し、棒の長短によつて數字の比較を行はしめんとする方法で、極めて廣く利用されてゐる。絕對數を示すときもあり、比率を示すときもあり、また其の各〻の場合に、構造狀態を示すときもある。上掲の假設例について見られたい。IはA、B、C三港の輸出價額を、IIは輸出價額と共にその內譯を、IIIは總價額中に占める製造品の百分比を、IVは內譯の百分比を示すものである。

棒圖表作製上の二三の注意を擧げよう。（一）右例では棒を縱に並べたが、勿論橫に並べてもよい。外國では文字が橫並べである爲め、橫の棒圖表が便利だと言はれてゐるが、日本では寧ろ縱の方が便利のやうである。（二）目盛の基準は零なるを原則とする。目盛が中途から始まつたのでは、高さの差は判つても、全體の長さは判らない。一〻

目盛に照し合はせれば別であるが、併し圖表は簡單にして正確な印象を與へるを目的とする以上、看者に餘計な苦勞を強要するが如き圖表は理想的とは言へない。零線を示す爲に餘りに棒が長くなつて却つて不體裁となる場合には、途中を省略してこれを齒狀の溝で示したり、或る目盛以下を省略して切目を齒狀とする。(三) 棒の幅は等しくする。元來棒圖表は棒の高さを利用したものであるから幅は無關係の如く思はれるが、併し異る幅の棒を並べるのは第一には體裁が惡く、第二には誤つた印象を與へる恐れもある。(四) 高さの順に竝べるのが、比較上最も便利である。

(2)面積圖表 (Area Diagram) 數字を圓又は正方形等の面積にて示した圖表をいふ。いまAB兩市の人口がそれぞれ五十萬人と二百萬人のとき、第二圖の如く示される。五十萬對二百萬であるから、一對四の面積にすればよい。正方形の面積は一邊の平方であるから、Aの一邊を一寸とすれば、Bの一邊は$\sqrt{4}$寸即ち二寸とすべく(I)、又、圓の面積は半徑の平方と圓周率(π)との相乘積であるから、



第二圖

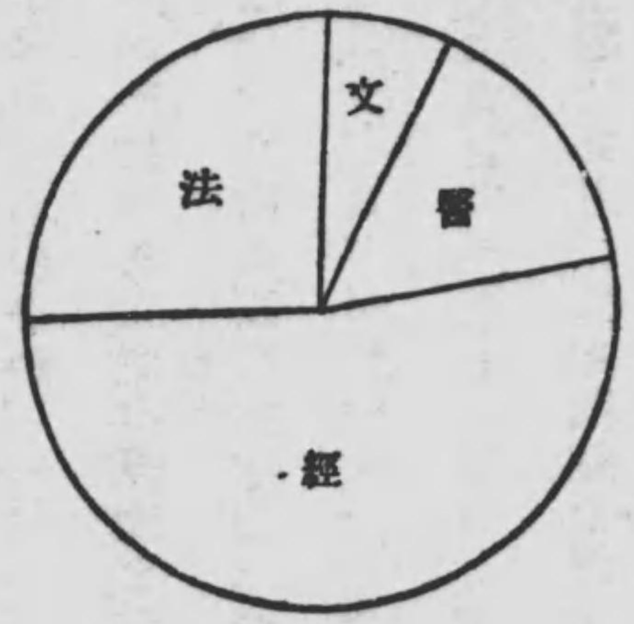
Aの半徑を一寸とすれば、Bの半徑は同じく$\sqrt{4}$寸即ち二寸とすべきである(II)。——(πは常數であるからこの場合關係はない)。即ち面積圖表を作るには先づ二つの數の比を求め(換言すれば一方は他方の何倍に當るかを求め)、次にこの比の平方根によつて作圖せねばならぬ。故に複雜な數字が與へられたときには可成り面倒な計算を必要とする。例へば甲乙兩市の人々が 512,532人 と 2,353,215人 だとすれば、先づ計算によつて 512,532 : 2,353,215=1 : 4.59 なるを求め、次に$\sqrt{4.59}$を計算する事によつて、兩者の邊の長さ(又は半徑の長さ)は 1 : 2.14 なる事が判るのである。

斯く計算が煩雜なるのみか、面積の比較なるものは一見して正確に行ひ難いものである。前圖を見て直ちに一對四の大さだと判る人は滅多に無いであらう。面積の下に數字を揭げて置けばよいが、併し勞して斯かる圖表を作るよりは、平易な棒圖表を選ぶ方が遙かに賢明である。

併し乍ら面積圖表の一種たるパイ圖表 (Pie Chart) なるものは、合理的でもあり且つ廣く用ひられてもゐるから、次に簡單に說明して置かう。この圖表は系列の內部構成比を示すを目的とするもので、その形がパイといふ菓子に似てゐるところからこの名稱がある。本大學は四學部から成り、學部の異るに從つて學生數も異る(上表參照)。いま學生總數に對する各學部の學生數の比率を求めれば第三欄の「%」の數字が得られる(これを求めるには例へば文學部に就ては $\frac{200}{3114}\times100 = 6.4\%$ とするのである)。この構成比に從つて一つの圓を分割すれば、一見して構成狀態を知り得るであらう。分割するには各構成比の割合に中心角(即ち三百六十度)を決定すべきであつて、1%は中心角3°6′に當るから、各構成比に3°6′を乘ずればよい。即ち文學部は 6.4×3°6′ = 23°。他の各學部も同樣にして求められる(第四欄の「中心角度」參照)。中心角度が判れば、分度器を用ひて容易にパイ圖表を描き得る。

慶應義塾大學
學部別本科學生數

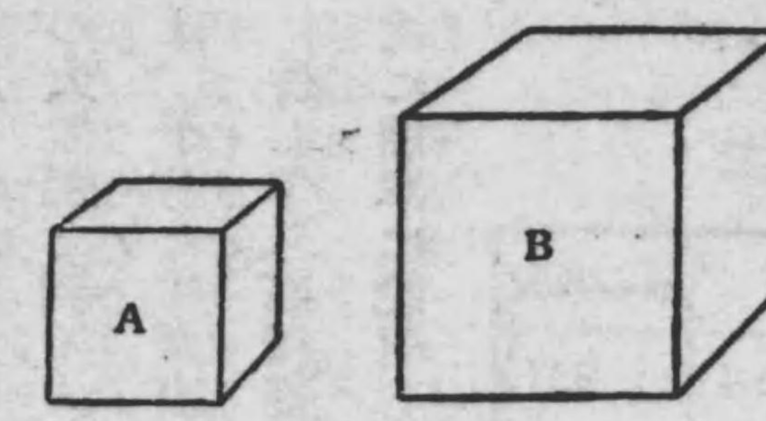
學部	學生數	%	中心角度
文學部	200	6.4	23°
經濟學部	1661	53.3	192°
法學部	798	25.7	92°
醫學部	455	14.6	53°
計	3114	100.0	360°



第三圖

パイ圖の長所は、單に弧の長さだけから數字を讀取り得るに在る。これは圓の部分の面積はその弧の長さに正比例するからで、この點で幅と高さを考慮せねばならぬ正方形又は圓の圖表とは趣きを異にするのである。即ち形は面積圖表ではあるが、實質的には寧ろ棒圖表に近いのである。

(3)體積圖表 (Solid Diagram)　前述の面積圖表が好ましくない理由は、面積なるものが常に幅と高さなる二つのディメンジョンから成る爲に、作製が困難なるのみか、一見して正確な印象を與へ得ない事である。然らば、これに更に「深さ」なる第三のディメンジョンを加へた立體の體積を以て數字を表はさんとすれば、作製も與へる印象も一層困難に一層不正確にならざるを得ない。面積圖表に例示した一對四の比を有するAB兩市の人口を立方體で示せば次の第四圖の如くなる。

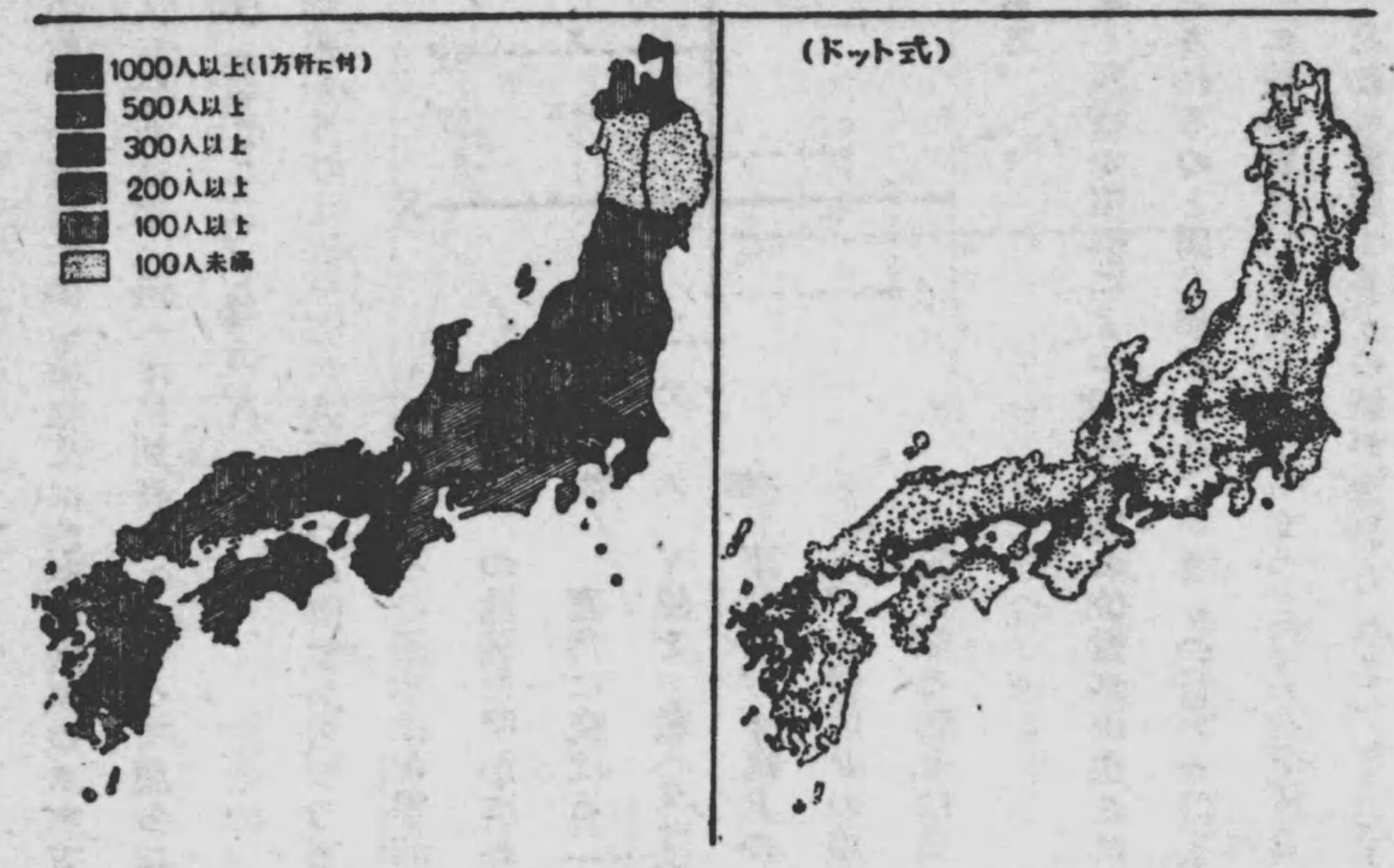


第四圖

これを描く爲には、A市の體積を一とせば(一邊の長さは勿論一になる)、B市の體積を四にするには一邊の長さを$\sqrt[3]{4}$即ち1.587とせねばならぬ。立方根を求めるのは、平方根を求めるよりも遙かに困難であり、而も困難を忍んで作つたABの二立體圖が、果して見る者に一對四の比と見て貰へるかどうかは可成り疑問であらう。

(4)繪畫圖表 (Pictorial Diagram)　日米兩國の軍艦噸數が夫々七十萬噸と百二十萬噸だとして、これを圖示するのに、七對十二の大小二隻の軍艦を描き、小で日本を、大で米國を示すやうな方法を繪畫圖表といふ。最も率直な印象を與へるから、展覽會などに於て盛んに見受けられるが、併し繪畫とは元來極めて複雜な立體圖であるから、正確に數字的比率を保つやうな大小の二圖を描くが如きは全く不可能といつてよい。繪畫圖表は元々通俗的用途にしか用ひられないから、概して數字

本邦人口密度



第五圖

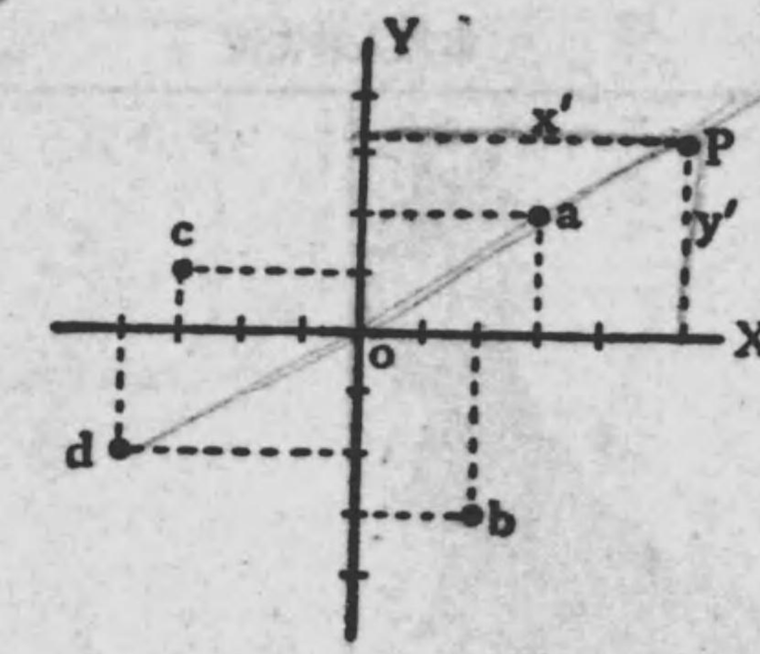
の差を誇張する繪畫が多い。せめてその下にでも實際の數字を掲げて置く位ゐの親切が欲しい。比較的無難な繪畫圖表は、例へば前例に於て日本は七隻、米國は十二隻の同一大さの軍艦を描く方法である。この場合には繪畫の大さの代りにその數が比較されるのであるから印象は極めて正確になる。通俗的用途には充分推賞し得る方法といへる。

(5)統計地圖 (Map Graph)　場所的系列の表現法として種種なる長所を有する圖表である。場所的系列とは、地域的分類による系列であるから、地圖の上に直接に何等かの樣式で數字を示せば、最も的確な印象を與へるであらう。例へば府縣別人口密度を地圖の上に大小の圓とか濃淡の色や模樣で示したり、又はドット式 (Dot Chart) といふ點圖で示したりする事は諸君も御承知の事と思はれる。統計地圖は普通の地圖と異つて、地圖そのものは格別正確なるを必要としない。大體の形狀が判れば充分である。猶ほ上の第五圖で判る通り、ドット式に據れば一府縣內部の人口分布狀態をも示し得て、

この點、一府縣は同一模樣又は色彩で塗りつぶす方法よりも優れてゐると言へる。尤もこれが爲には豫め各府縣內部の人口分布狀態（例へば郡別密度）が判つて居らねばならぬ事は言ふ迄もない。

(B) 變數が二つの場合

時系列及び度數分布が共に二つの變數から成る事は既に述べた。即ちこれらは數と數との關係を示すものであつて、その圖示は數學的の線によるを原則とする。その說明に入るに先だち、豫め數學線の性質を明かにして置かねばならぬ。先づ點の圖示から始めよう。

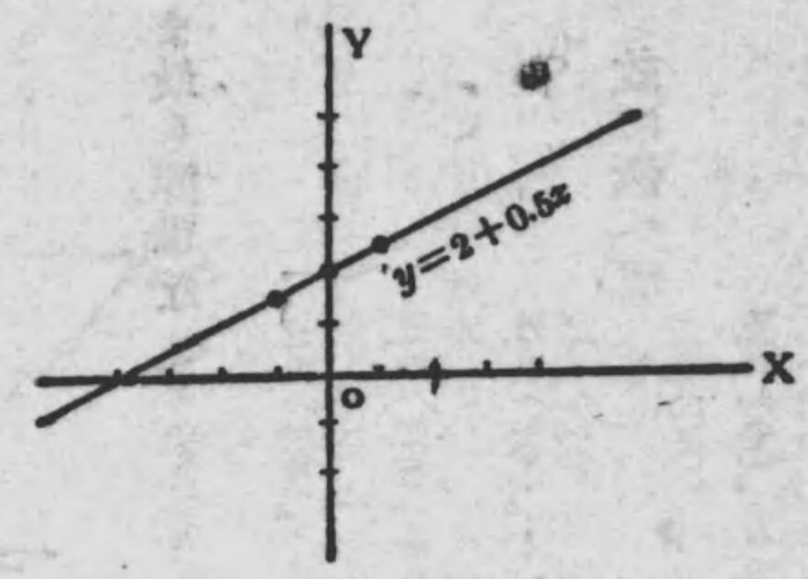


第六圖

直角に交はる二直線X及びYをとり、交點をOとする。X線をx軸（又は橫軸）、Y線をy軸（又は縱軸）、Oを原點といふ。いま定點Pから兩軸に垂直線x'及びy'を下せば、x'はPのx座標、y'はPのy座標なりといひ、この二つをP點の座標といふ。一般にPの座標を、P（x, y）と記す。x座標はy軸から右に測るときは正、左に測るときは負であり、y座標はx軸から上に測るときは正、下に測るときは負である。

斯く縱橫の座標によつて或る點の位置が決定される事は、恰も地球表面上の任意の一地點が經度と緯度とによつて決定されるのと同じ事である。いま a(3,2), b(2,-3), c(-3,1), d(-4,-2) なる四點を圖示すれば上掲、第六圖となる。

次に線の數學的性質を考察しよう。$y=2+0.5x$ なる x に關する一次式が與へられたとする。x に種々の値を與へれば、それに對應する y の値が求められる。$x=0$ とすれば $y=2$ であり、$x=1$ とすれば $y=2.5$ である。更に $x=-1$ とすれば $y=1.5$ となる。（x の變化に從つて y が變化するのであるから、x を自變數、y を屬變數といふ。）これらの x、y の一對の値は、前述の點の圖示法によつて圖表の上に表はす事が出來る。

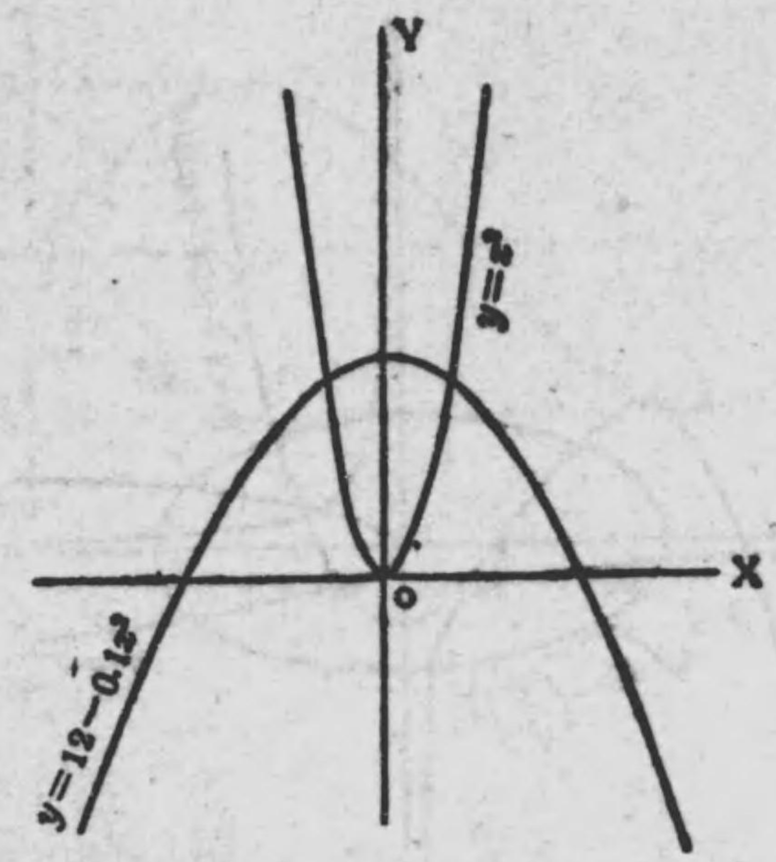


第七圖

これらの點を連結すれば第七圖の直線が得られる。總て $y=a+bx$ なる x に關する一次方程式は斯く直線となるのであつて、目盛に對照すれば、x に對應する y の値は容易に求められるのである。

同樣に x に關する二次方程式 $y=a+bx+cx^2$（上の一般式に於て、$a=0$, $b=0$, $c=1$ としたもの）につき、x に種々なる値を與へれば、その都度それに對應する y の値が得られる。即ち

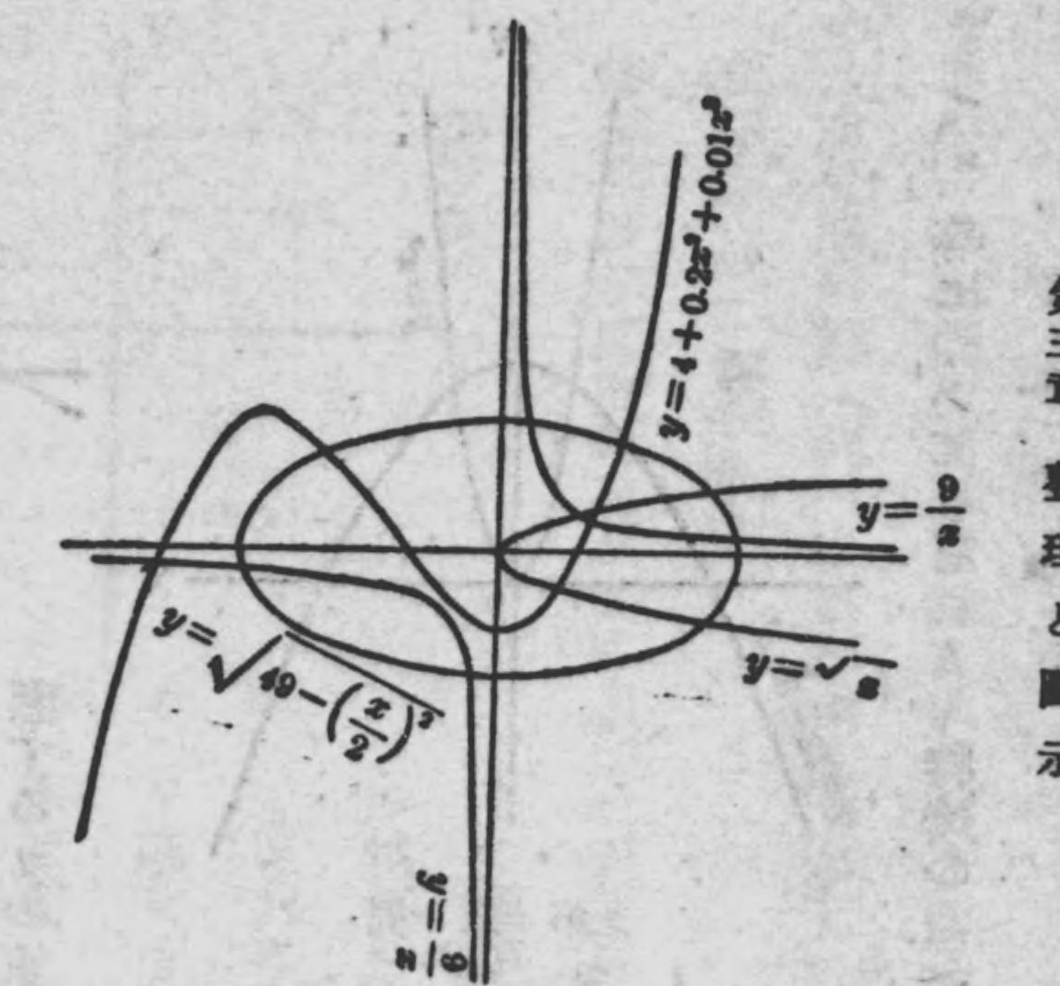
x	-4	-3	-2	-1	0	1	2	3	4	5……100
y	16	9	4	1	0	1	4	9	16	25……10000



第八圖

これ等の諸點を座標點にて示し、相互に連結すれば上圖の如き曲線となる。斯かる曲線を拋物線といふのである。同じく $y=12-0.1x^2$ も亦拋物線である。一次式が直線となるに對し、二次式は總て拋物線の形を採るのである（第八圖）。

更に三次式 $y=a+bx+cx^2+dx^3$、有理分數函數 $y=\frac{a}{x}$ 又は無理函數 $y=\sqrt{x}$ 等についてもそれぞれ特殊の曲線が求められる。これらは一括して次の第九圖に揭げて置いた（P. Lorenz



第九圖

Höhere Mathematik より採る)。

最後に超越函數の一たる三角函數の圖示を説明せねばならぬ。三角函數は一定の周期と振幅とを以て規則正しく變動しつゝある循環的運動を示す爲に用ひられる。自然現象にしてこの種の運動を示すものは固より多いが、社會現象、就中經濟現象の中にも決して少くはないのである。季節變動は勿論、景氣變動も亦ある程度の規則的循環を示すが故に、時系列の解析に於て三角函數の知識は絶對的に必要と言つてよい。その詳細は後章に讓るとして、こゝでは一般に三角函數が如何なる曲線によつて示されるかを説明して置かう。いま最も簡單な $y=9\sin x$ なる函數を想定する。三角法の原理に從へば

$$\sin(180°-x)=\sin x$$

$$\sin(180°+x)=-\sin x$$

$$\sin(360°-x)=-\sin x$$

故に例へば

$$\sin 160°=\sin(180°-20°)=\sin 20°=0.3420$$

$$\sin 200°=\sin(180°+20°)=-\sin 20°=-0.3420$$

$$\sin 340°=\sin(360°-20°)=-\sin 20°=-0.3420$$

又 x が 360° 以上の角度に於ては $\sin x$ の値は、x から 360° 又はその倍數を差引いたものに等しいから

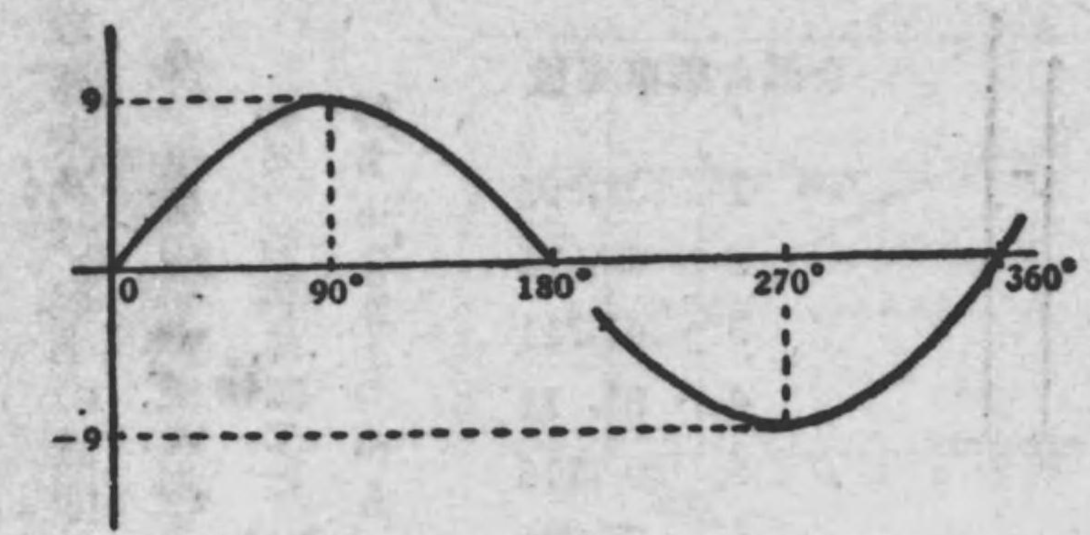
$$\sin 380°=\sin(380°-360°)=\sin 20°=0.3420$$

$$\sin 920°=\sin(920°-720°)=\sin 200°=-0.3420$$

この原理を應用して上記の函數 $y=9\sin x$ の x に種々なる値を與へ、それに對應する y の値を求めれば次の如くであり、それを圖示すれば第十一圖となる。

x	0°	45°	90°	135°	180°	195°
$\sin x$	0.0000	0.7071	1.0000	0.7071	0.0000	−0.2588
y	0.00	6.36	9.00	6.36	0.00	−2.33

x	225°	270	300°	360°	405°
$\sin x$	−7.7071	−1.0000	−8.660	0.0000	0.7071
y	−6.36	−9.00	−7.79	0.00	6.36



第十一圖

この圖に於て0°から360°までで一つの波動が完了し、360°以上は前の原理に從つて全く同一の波動が繰返へされるのであるから、極めて規則正しい循環なる事が判らう。この波動に於ては、0°から360°までの長さを周期といひ、これに對して其の間に於て y のとり得

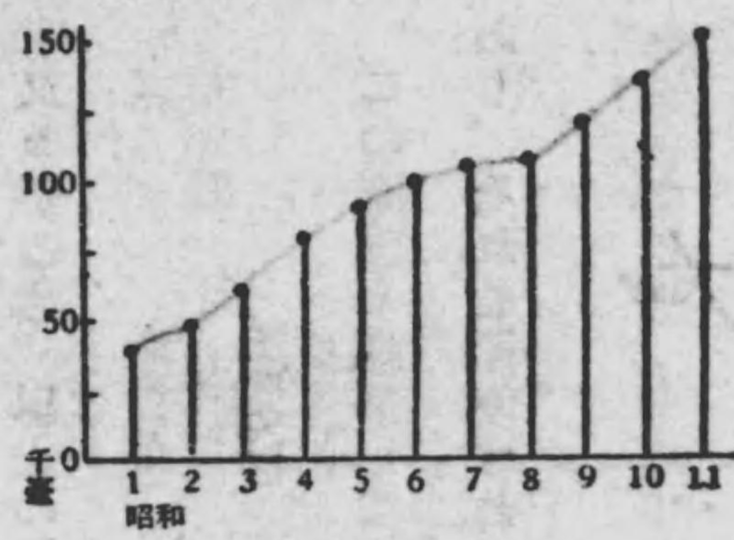
る極大及び極小の値即ち9から−9までを振幅といふのである。(註)

註　振幅はyの極大値及び極小値によつて決定される。然るにかゝる値は、與へられた函數の微係數を零と置く事によつて求められる。右例に於ては$\frac{dy}{dx}=9\cos x$であるから、$9\cos x=0$としてxの値を求め、その値を始めの式に代入してyを求めればよい。即ち$9\cos x=0$から、$x=90°$となり、これを$y=9\sin x$に代入すれば$y=9\sin 90°=9$が得られるのである。

全國自動車臺數

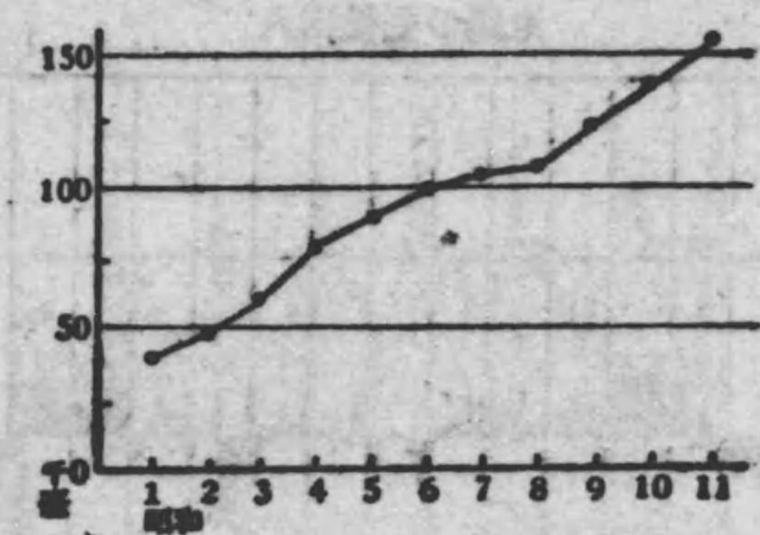
昭和	1	38,693
	2	49,675
	3	61,711
	4	81,[illegible]71
	5	9[illegible],116
	6	98,996
	7	103,905
	8	106,783
	9	121,192
	10	134,859
	11	149,635

(a) 時系列の圖示　上の如き昭和元年から十一年に至る十一年間の全國自動車臺數の系列が與へられたとする。各年の自動車臺數をy軸に、年度をx軸にとり、――時系列の圖示に於ては年度を常にx軸にとる。これは年度なるものは自變數だからである――各年の臺數を高さで示せば、上掲第十二圖を得る。併し時なるものは元來無始無窮の連續性を有し、從つて時と函數關係に立つところの臺數も亦連續性を帶びる事になる。從つて昭和元年の三八、六九三臺は翌二年の四九、六七五臺に、後者は同じく翌三年の六一、七一一臺に連結してゐるのである。斯くて圖示に際してもこの連續性を示して置くのが合理的である。これが爲には第十二圖の各垂線の頂點を順時に結べばよい(第十三圖)。これが時系列圖示に於て最も一般に用ひられる方法である。

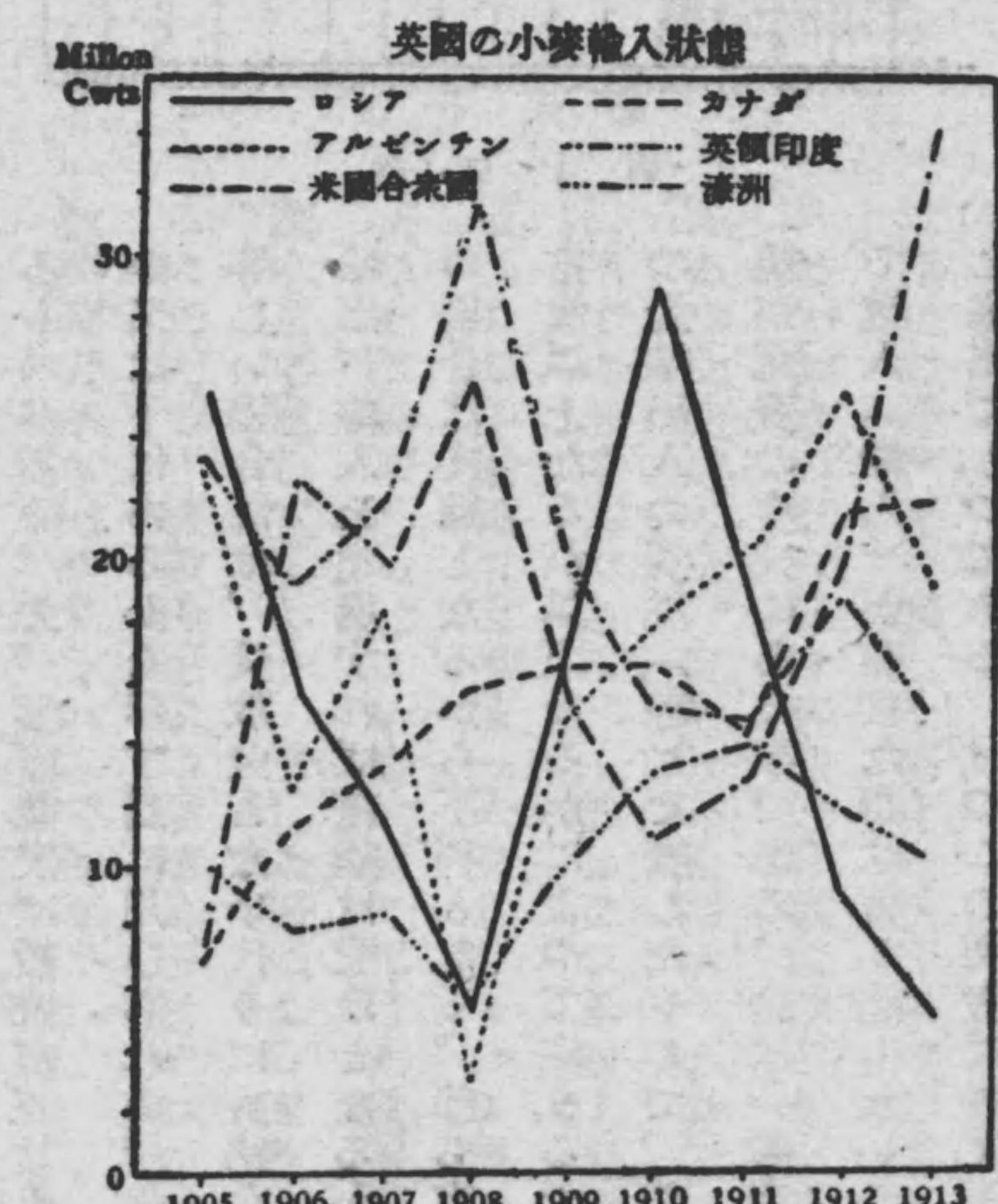


第十二圖

併し連續の概念を一層徹底せしめるならば、數箇の直線から成る多角形狀の方法は未だ不充分なる事が判る。臺數の增減は漸時に行はれるのであるから、線に角のある可きものではない。若し右の如き年別統計の代り、より詳しい月別統計が與へられてゐたとすれば、右と同樣の方法によつて求められた多角形は、非常に多くの角を有する事となり(前例で言へば11×12箇月即ち百三十二角となる)、寧ろ曲線に接近するのである。故に或る操作を施して滑かな曲線と化して了ふのが最も合理的といはねばならぬ。併しその方法は線の當嵌め(Curve-fitting)なる特殊の技術を必要とする。この技術は時系列解析の主要問題であるから、後章に改めて説明する。



第十三圖



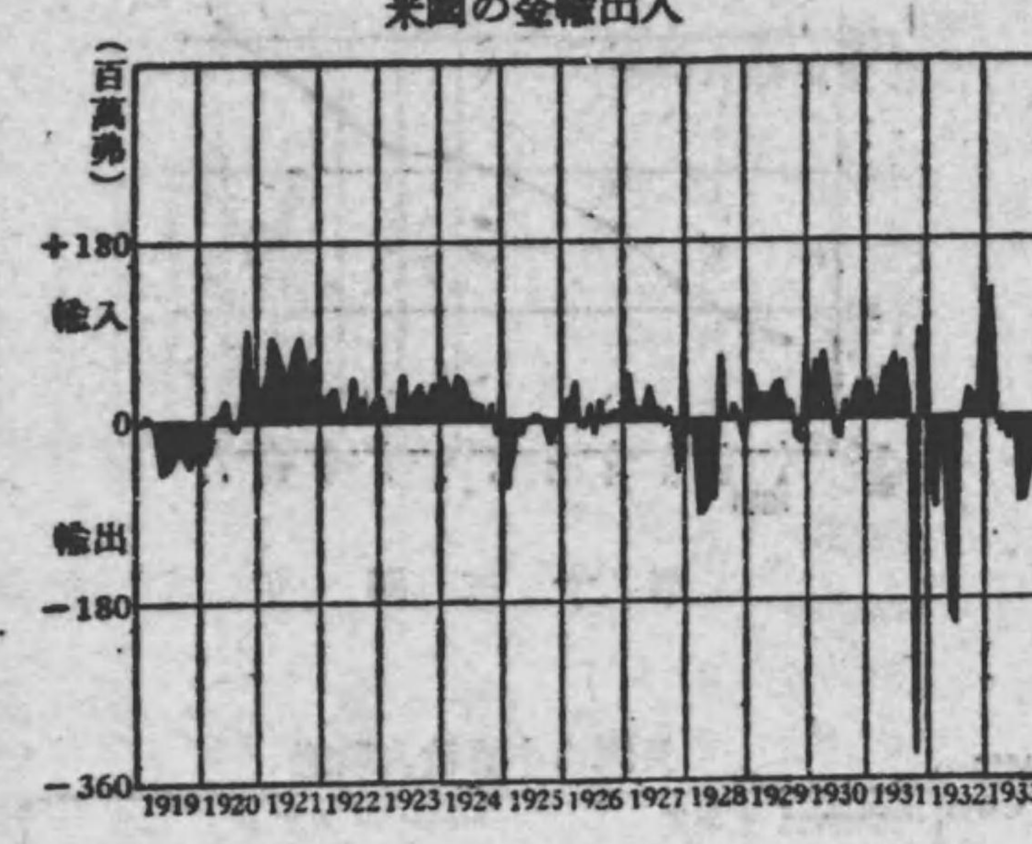
第十四圖

次に若干の實例を擧げて參考に供しよう。第一に、同一圖表に數箇の線を示さうとする場合には、判別を容易ならしめるやう多少の工夫を必要とする。異る色を用ふれば問題はないが、同一色で示す場合には、實線、點線又は波線等を巧みに使ひ分けねばならぬ(第十四圖參照)。線が二つしか無い場合には、一方を太い實線、他方を細い實線で示すのが一般であるが、併し線なるものは元來理論的には長さはあつても幅はない筈のものであるから、餘りに太い線は不合理

であり、且つ目盛との對照が不可能となる。加之、幅に大小のある二線は、恰もその對象の重要さに輕重あるかの如き印象を與へる恐れもある。線の種類や太さについては作製者の常識に俟つ外はない。

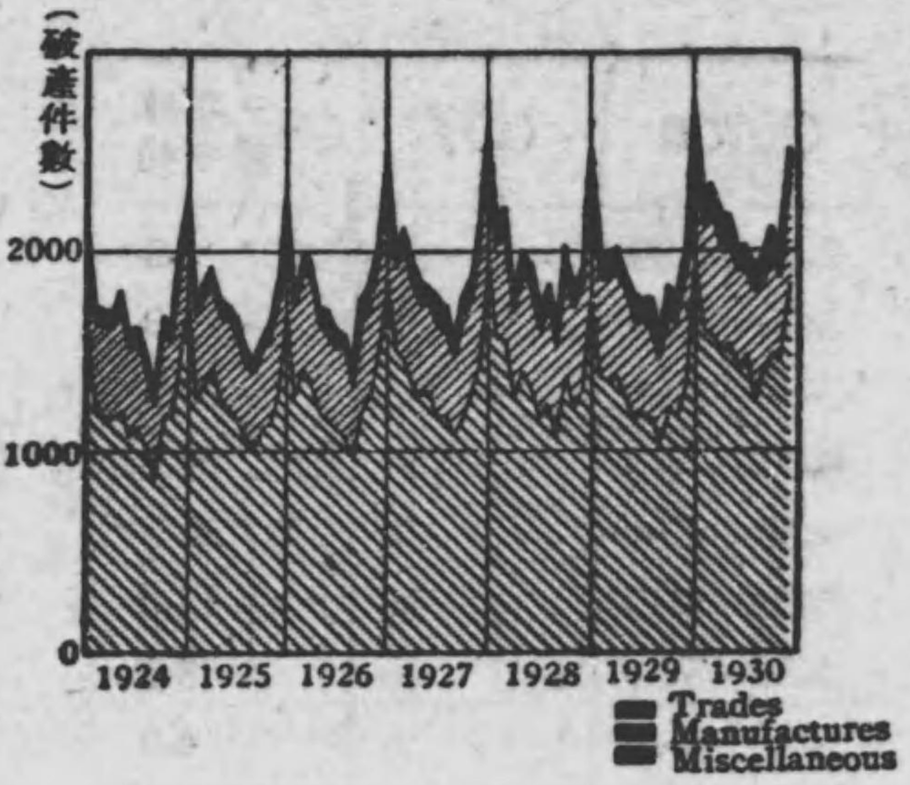
線圖表に面積圖表の意義を加味せしめたものに、影繪圖表（Silhouette Chart）及び帶狀圖表（Band Chart）なるものがある。先づ影繪圖表は數値が正常値（normal value）の上にあるか又は下にあるかを示す爲に用ひられる。例へば貿易は輸出と輸入とが相等しい場合には、受取金もない代りに支拂金もなく、即ち正常狀態である。輸出が輸入を超過すれば差額は受取となるから有利となり、反之、その逆の場合は支拂となるから不利となる。即ち有利とか不利とかは、それが正常値以上なるか以下なるかによつて定まるのである。上の第十五圖は米國の金輸出入の帳尻（差）を表したもので、普通の圖表と相違する點は、零線が圖表の中央にある事である。金は一般商品と異り流出（輸出）が不利で流入（輸入）が有利な事は言ふ迄もないから、零線以上を輸入に、以下を輸出にとつてある。且つこの圖表に於て吾々の知らんとするものは、線の動きよりは、寧ろ線によつて決定される面積なのである。何となればこの面積によつて、一定期間内に受取が多かつたか（零線以上の面積の方が大きかつた場合が之である）、又は反對に支拂が多かつたかが判るからである。斯くてこの上下の面積を明示する爲に、線と零線の間を黒一色に塗りつぶして置けば、甚だ判別が容易となるのである。

米國の金輸出入



第十五圖

第二の帶狀圖表は總數及びその構成部分の變化を示す爲に用ひられるものである。これに就ては別に說明を要しないと思はれる。次の第十六圖について知られたい。これは米國の年別及び職業別の破産數を示すもので、既に說明した構成棒圖表に連續の概念を加味したに過ぎない。



第十六圖

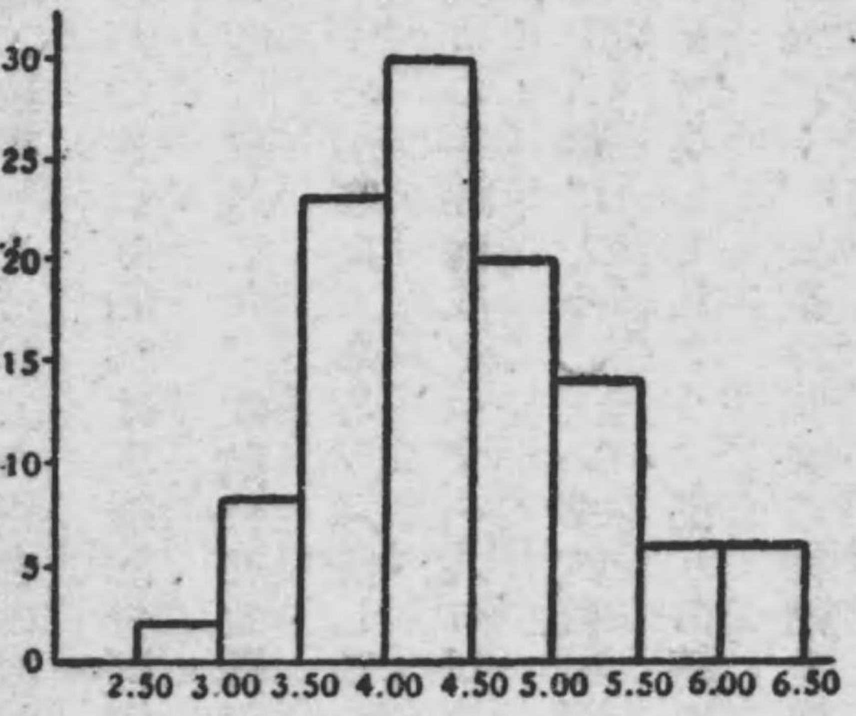
(b) 度數分布の圖示　度數分布とは既に說明した通り、量的屬性（年齡・身長・賃銀額等）によつて統計量（人口又は職工數等）を分類した場合の統計系列である。時系列又は場所的系列に於けるが如き自然的又は客觀的分類標準がない爲め、豫め階級を定める必要のある事は前述した。度數分布を圖示する場合には、階級をX軸に度數をY軸に示すのが原則である。（併しこの原則は特殊の事例では通用しない。例へば人口の年齡構成を示す「人口ピラミット圖」に於ては、年齡階級をY軸に、人口をX軸にとる。）いま次頁に掲げた賃銀別職工數の統計を用ひて、その圖示を考察して見よう。但し曲線による度數分布圖表即ち度數分布曲線に入るに先だち、度數棒圖表の構造を理解せねばならぬ。蓋し度數分布曲線は度數棒圖表から出發するからである。

職工數を分類した結果は、次表の第一及び第二欄によつて示される度數分布表となる（第三欄は暫く措く）。階級の大さは五十錢であるから、X軸を五十錢の幅に目盛を打つ。但し最小限界が二圓五十錢であるから、零から二圓五十錢迄の幅は全く省略するか又は短縮して差支へない。いま各階級に含まれる職工數（即ち第二欄のf）をそれぞれ

棒の高さで示せば前述の棒圖表に似た次の第十七圖が得られる。これを度數棒圖表又はヒストグラム（Histogram）といふ。一般の棒圖表と異る點は、X軸が目盛を打たれてゐること、從つて棒の位置は目盛（即ち階級の順序）によつて規定され、棒圖表の如く任意にその順序を變へる事は許されない事である（棒圖表では、なるべく高さの順に並列せしむべき事は既に述べた）。いま階級を示すX軸の目盛を見るに、元來それは一定範圍の賃銀（右例では二圓五十錢から六圓五十錢まで）を便宜上五十錢づつの大さに區劃したもので、從つてこれら區劃は各自獨立せるものではなく、實は一箇の連續的なものなのである。即ちその性質は、時系列圖示に於けるX軸（時）と殆ど同種のものである。こゝに「殆ど」と言つたのは、實は度數分布に於ける連續性は、必ずしも時系列に於けるが如き嚴格なものではないからである。時は無限に細分し得るに反し、例へば室數による家屋の分類に於ては一室を以て最小限度とし、一室の家屋の次には二室の家屋が來るのであつて、一・二室とか一・八室とかの家屋は考へられない。即ち、この種の系列は形式上は度數分布であるが、實質的には寧ろ質的屬性系列と言つてもよい。併し賃銀の如きは額に比して單位が非常に小さいから殆ど連續的と見てよいのである。殊

(1)賃銀	(2) f	(3)三項移動平均
2.50—2.99	2	3.3
3.00—3.49	8	11.0
3.50—3.99	23	20.3
4.00—4.49	30	24.3
4.50—4.99	20	21.0
5.00—5.49	13	13.0
5.50—5.99	6	8.3
6.00—6.49	6	4.0

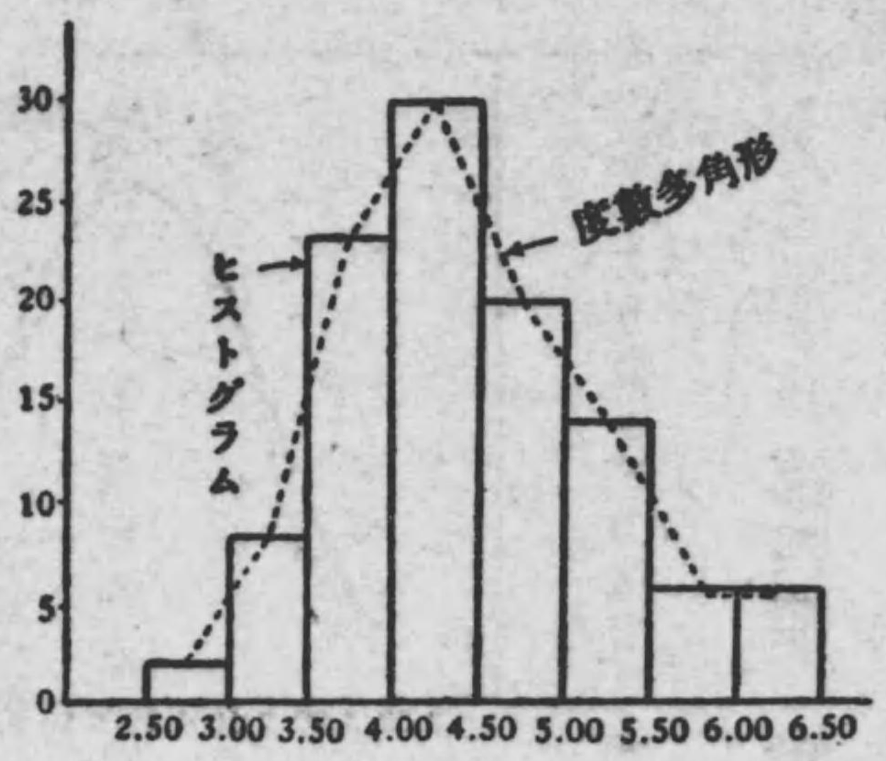
度數棒圖表
（ヒストグラム）



第十七圖

に例へば七十圓の月給を日給に換算すれば（一月を三十日として）、二圓三十三錢三厘三毛三三三……となり、室數の場合とは著しく趣きを異にする事が判らう。

故に斯かる連續性を考慮に入れゝば、時系列の場合と同じく、棒の高さは相互に連結されねばならぬ。棒は一定の幅を有するから、上邊の中央點を結ぶ。例へば四圓五十錢から五圓までの階級に於て上邊の中央點をとれば、そのx座標は四圓七十五錢となる。これは四圓五十錢と五圓との平均に外ならず、從つてその意味は、右階級に含まれる三十名の職工は、平均四圓七十五錢の賃銀を受けてゐるといふ事である。各階級についてその中點を求め、それらを相互に連結したものを度數多角形（Frequency Polygon）といふ（第十八圖）。そして更に連續性の概念を一層徹底せしめるならば、時系列に於て説明した通り、斯かる度數多角形は曲線形に改められねばならぬ。これを度數曲線（Frequency Curve）といふ。

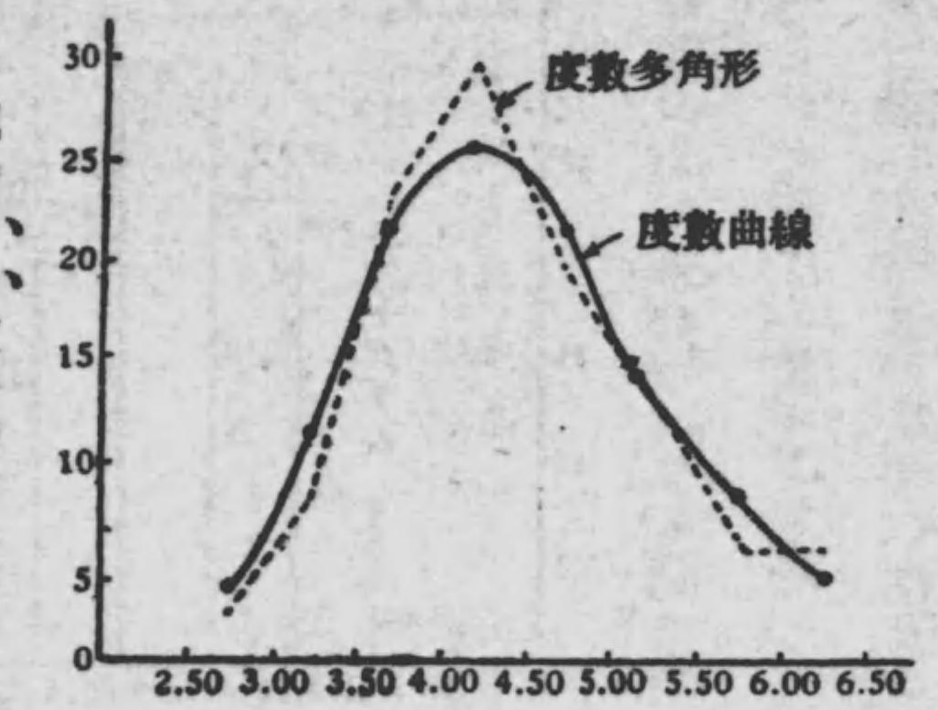


第十八圖

度數曲線を描くには目測によるか又は移動平均法なる特殊の平均法を用ふる。階級の間が狹い時には目測で角をならせばよい。目測に自信が無ければ移動平均を用ふる外はない。移動平均とは與へられた順序の數字を數項づつ順次に平均する事である。六二頁の表の第三欄は三項移動平均の値を示したもので、第二欄の度數fを最初には0 2 8を平均して3.3を求めてこれを第一の階級の度數とする。次には2 8 23の平均20.3を第二の階級の度數とし、以下同樣にして算出したのである。第一の階級の移動平均値を求める爲に零を採り入れたのは

11.0

最初の度數が小さかつたからで（この例では2）、一種の便宜手段である。最後の階級の移動平均値4も同様にして6　6　0の三數を平均したものである。移動平均は元來系列の不規則性を修正する爲のもので、從つてこれによつて求めた如上の各平均値は修正された度數と認められるのである。



第十九圖

斯くて與へられた度數の代りに、これら修正された度數の座標點を打ち、順次に滑かな曲線で結べば求むる度數曲線が得られる（第十九圖）。猶ほ幾何學的作圖法によつて右と同一の座標點を求める方法もあるが、此處には省略する。移動平均については後に適當な個所に於て改めて説明しよう。移動平均法による曲線化は時系列に於ても利用される事があるが、最も廣く用ひられるのは、最小自乘法その他のより數字的方法である。

第四節　對數圖表

對數が一般數學に於て如何に廣く利用されてゐるかは特に説く迄もないが、統計的研究に於てもその利用範圍は頗る廣いのである。その一例は時系列解析の中に出て來る季節指數の算出に見られる。故に對數圖表の説明に入る前に、簡單に對數そのものゝ意義を述べよう。尤もこゝで扱ふものは常用對數についてゞある。常用對數と並んで屢ゝ使用される自然對數は、初等統計學では省略してよいと考へる。

或る數を10の羃と考へたとき、その指數（この場合の指數 Exponent と、前述した比例數としての指數 Index-number とは別のものなる事を注意されたい）を指してその數の對數といふ。例へば、$100=10^2$ であるから、2は100の對數である、或ひは100の對數は2であるといひ、これを $\log 100=2$ と記す。同様にして $\log 1=0$, $\log 10=1$, $\log 1000=3$ であり、また $\log 0.1=-1$, $\log 0.01=-2$ である。さて10の對數は1で、100のそれは2であるから、10から100までの或る數、例へば50の對數は言ふ迄もなく1と2との間の或る數でなければならぬ。對數表によつてこれを求めれば 1.69897 である。書換へれば $\log 50=1.69897$ であり、その意味は $50=10^{1.69897}$ といふ事である。斯く如何なる數も10の羃と考へ得る事が、對數の根本概念である。

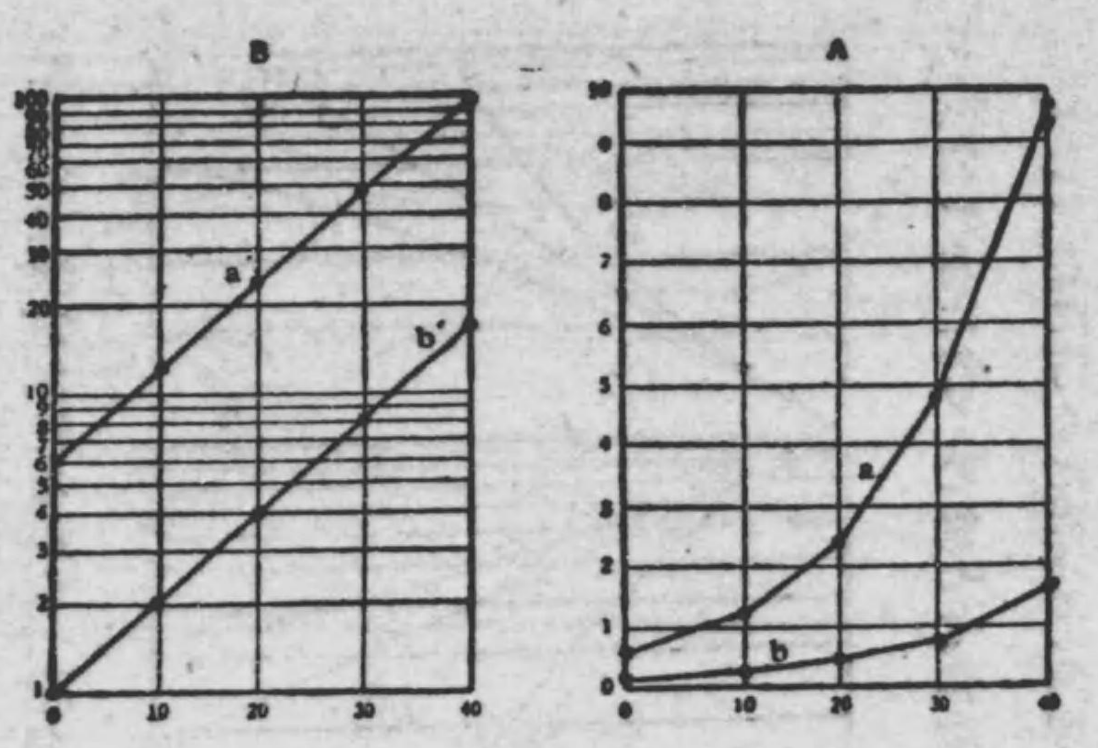
次に $A=10^a$, $B=10^b$ だとして、指數の法則によつて $A\times B=10^a\times 10^b=10^{a+b}$ であるから、兩邊の對數を求めれば、$\log(A\times B)=a+b=\log A+\log B$ である。同様に $\frac{A}{B}=\frac{10^a}{10^b}=10^{a-b}$ であるから、$\log\left(\frac{A}{B}\right)=a-b=\log A-\log B$ となる。或る數の對數がいくらかといふ事は、對數表から直ちに求められるし、また對數が與へられゝばその眞數（前例で云へば 1.69897 なる對數の眞數は50である）がいくらであるかも、對數表を逆にひけば簡單に求められる。そして上記の如く、對數を用ふれば、乘算は加算の、除算は減算の手續きで答を得られるから、複雜な計算もこれによつて至つて平易なものとなるのである。（註）

註　例へば $x=\frac{526\times 325^3}{450^2}$ を求めるに、兩邊の對數をとれば $\log x=\log 526+3\log 325-2\log 450$ となり、對數表から $\log x=2.7209857+3\times 2.5118834-2\times 2.6532125=4.9502109$ これは求むる數 x の對數であるから、その眞數卽ち x は、同じく對數表から $x≒89168$ なる事が判る。（詳しくは $x=89{,}168.387755097$）

但し注意すべきは、10を何乘するも零又は負數とはならぬから、零又は負數の對數は無い事である。斯かる數が算式の中に介入してゐるときは對數による計算は出來ないのである。

對數圖表とは數字そのまゝの大さの代りにその對數の割合に目盛を刻んだ用紙（對數用紙）に描いた圖表である。嚴密な意味に於ける對數用紙は縱橫ともに對數目盛となつてゐるが、統計グラフとして時系列の變化を示すを目的とする場合には、橫の目盛は普通の均分目盛で、縱の目盛だけが對數目盛になつてゐる所謂「半對數用紙」Semi-logarithmic Paper を使用する。蓋し橫目盛には時の經過を記すから、これに就ては均分目盛でなければ意味がないのである。即ち半對數用紙の縱目盛は、1. 2. 3. 4. 5. 6……等の對數の割合、即ち 0, 0.30103, 0.47712, 0.60205, 0.69897, 0.77815,…… の割合となつてゐる。そして1から10までの幅と、10から100、及び100から1000までの幅は各々等しい。蓋し $1\,0=10^1$, $100=10^2$, $1000=10^3$ であるから。換言すれば1が10にまで增加する割合は、10が100にまで、及び100が1000にまで增加する割合と同じく、何れも十倍であるから。同樣にこの目盛に於ては、等しい倍數に當る間隔は常に同一である。例へば1から2、2から4、4から8、8から16の幅は同一であり、また1から3、3から9、9から27の幅は同一である。

斯くて普通の方眼紙の上では1に1を加へるも、10に1を加へるも、その增加の幅は同じであるが、對數用紙の上では1に1を加へる事は10に1を加へる場合の十倍の幅となつて表はれる。いま同じ數列を普通目盛に記入した場合と半對數目盛に記入した場合との相違を圖示しよう。即ちこゝに十年で元利合計が倍加する利率で一定金額が預金されたとき、四十年間の增加狀態を示めさうとする場合、普通の目盛では始めの金額が大きいか小さいかによつて全く異つた線が描かれるが、半對數目盛では始めの金額如何に拘らず同一形の線が描かれる（第二十圖參照）。



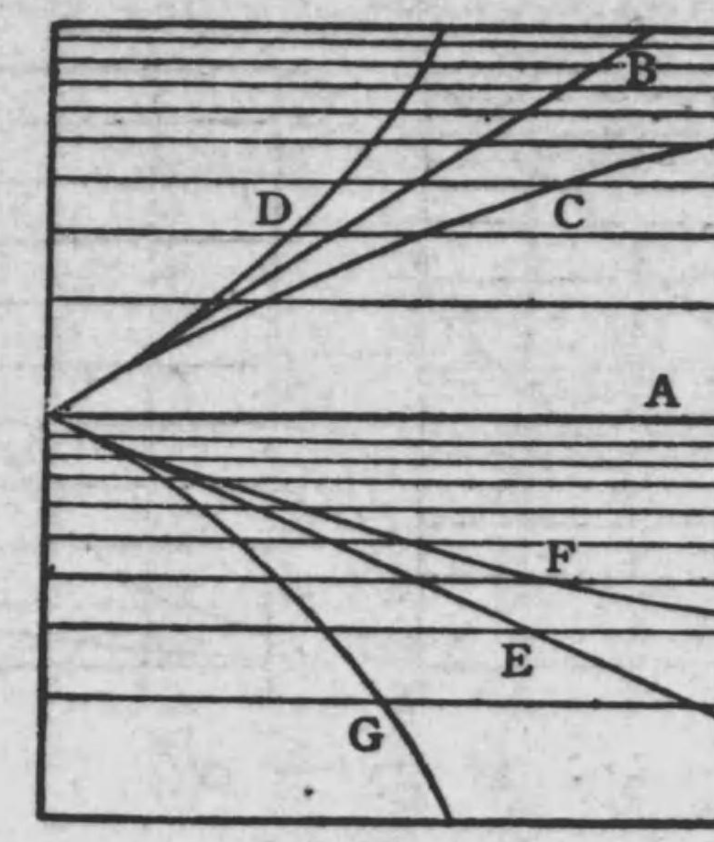
第二十圖

この圖では最初の金額は六圓と一圓である。A表即ち普通の目盛に於ては、a及びbは共に屈折線で且つ全く異つた形を示す。これは最初の金額六圓が四十年後に九十六圓になるのと、一圓が四十年後に十六圓となるのとの間には、絕對額に大なる相違があるからである。反之、B表即ち半對數目盛に於てはa′とb′は共に直線で且つ全く平行となる。何故直線となるかと言へば、兩者共に全期間を通じて同じ割合で增加してゆくからであり、また何故平行線となるかと言へば、兩者の增加率が等しいからである（兩者ともに十年にして元利合計を倍にするなら、同じ利率即ち同じ增加率である事は言ふ迄もない）。一般に複利の公式 $y=p(1+r)^x$（pは元金、rは利率、xは年數を示す）の兩邊の對數をとれば、$\log y=\log p+x\log(1+r)$ となり、これは一次方程式であるから、對數目盛の上では當然直線となるのである。そしてこの直線の傾きの角度は r（利率）のみによつて決定されるから、p（元金）の大小は毫も關係がない。a′b′が平行線となるのもこれが爲なのである。

半對數用紙そのものは、對數の原理に據つて作られたものであるが、これを利用するには格別對數に關する知識を必要としない。普通の方眼紙と同樣に、單に目盛の數字に從つて記入して行けばよいのである。唯だ忘れてならぬことは、普通の目盛は0から出發するに對し、對數目盛は普通は1（特に小さい數を記入する場合は0.1又は0.01等）から

出發し、決して0から出發しない事である。これは前に述べた對數の性質から明かであらう。

要するに對數圖表の主たる目的は增減の割合を簡單且つ的確に示すに在る。併し目盛と照し合はせれば絶對量を讀む事も可能であるから、算術圖表の役目をも兼ねてゐると言へる。對數圖表に示される線の角度は增加率又は減少率によつて決定されるから、圖表を一瞥して直ちに增減の比率を推知し得るのである。即ち第二十一圖の示すが如く、



第二十一圖

（一）線が底線に平行なる時は量に增減なきを示し（上圖のA線）

（二）上昇的直線は量が一定比率にて增加するを示し（B線）

（三）右に行くに從つて角度の減ずる上昇線は、量は增加しつゝあるも其の率は漸減しつゝあるを示し（C線）

（四）右に行くに從つて角度の增す上昇線は、量も增加率も共に漸增しつゝあるを示し（D線）

（五）下降的直線は量が一定比率にて減少しつゝあるを示し（E線）

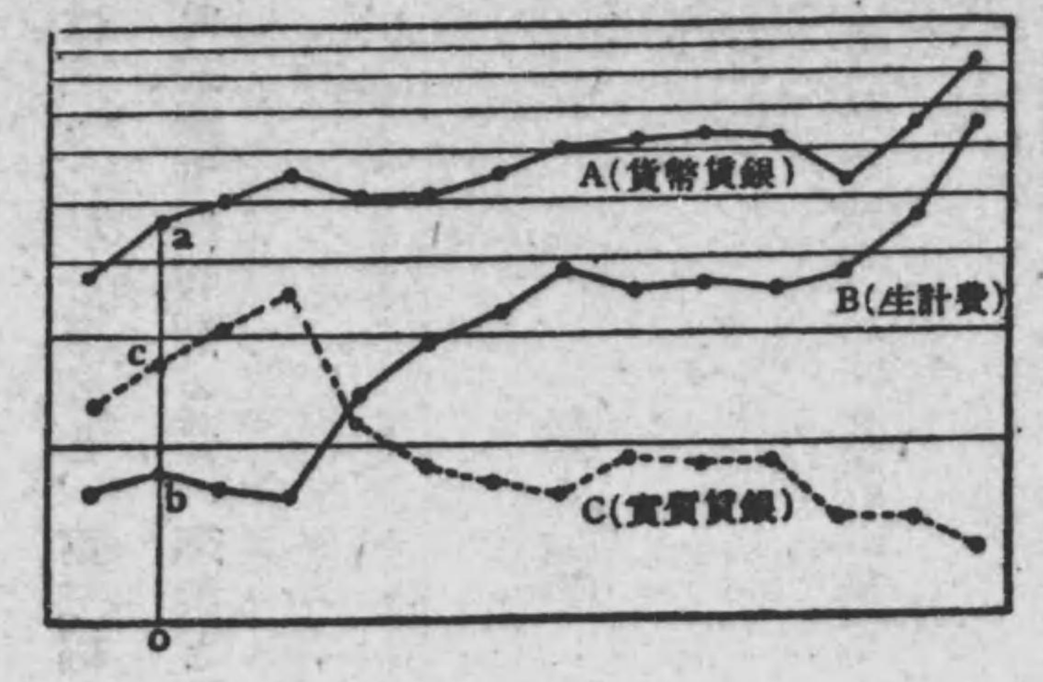
（六）右に行くに從つて角度の減ずる下降線は、量は減少しつゝあるも、減少率は漸減しつゝあるを示し（F線）

（七）右に行くに從つて角度の增す下降線は、量が減少すると共に減少率は漸增しつゝあるを示す（G線）

（八）同一圖表に示された二線が、平行なるときは增加又は減少の率が同一なるを示し、平行ならざるときは、急峻な線はより大なる增加率又は減少率を示す。

對數圖表の第二の用途は、同一圖表の上に、甚だ桁の違つた數列を描き得る事である。算術圖表では斯かる場合には極めて厖大な紙面を必要とし、總てに於て不便である。

第三の用途は、圖表の上で簡單に乘除算を施し得る事である。前に説明した通り、對數による計算では乘算は加算、除算は減算の形で行ひ得るから、對數圖表では、或る數に或る數を乘ずる必要ある場合には加へ合はせればよく、割る必要ある場合には差引けばよい事になる。いま勞働賃銀に就て見るに、貨幣賃銀の騰落は必ずしも勞働者の利不利と合致するものではない。貨幣賃銀が騰貴しても、若し物價の騰貴がより甚だしければ生活は苦しくなるし、反對に貨幣賃銀が下落しても、若し物價の下落がより甚だしければ、生活は却つて樂になる。即ち生活問題として賃銀を考察する場合には、貨幣賃銀と一般生計費との關係を見ねばならぬ。これが爲には貨幣賃銀をその時の生計費指數（これは後に説明する）で割る事によつて實質賃銀に換算すればよいのである。從つて對數圖表の上に貨幣賃銀と生計費指數とを示す二線を描き、各時點に於て、兩者の高さの差を記入すれば、求める實質賃銀が得られるのである。



第二十三圖

上圖の假想例に於て、$co=ao-bo$ である。換言すればbからaまでの距離をコムパスで測り、その長さをoに立てればcが得られる。これを各時點について行ひ、得た諸點を結べばc線、即ち實質賃銀の變動が表はれて來るのである。

註　圖示法を通俗的に説明したものとして次の二つは便利である。
猪間驥一――經濟圖表の見方書き方使ひ方（東洋經濟新報社）

白崎享一――統計グラフの書き方（圖勢社）

×　×　×　×

以上説明した統計學史、調査、分類及び圖示法の如きものは何れも統計的解析に入る準備階段である。次の第二編は箇々の統計系列を、第三編はこれら統計系列間の關係を主題とする。即ち第二編では度數分布及び時系列の解析を、第三編では相關々係の理論と測定を述べたいと思ふ。

第四章　度數分布の解析

量的屬性に基く分類の結果、即ち、度數分布なる統計系列は、度數が各階級に如何に配分されてゐるかを、問題とする。

これが解析は、具體的に言へば、(1)度數分布の形態、(2)平均値、(3)散布度、(4)非對稱度、の四つを明かにしなければならぬ。

第一節　度數分布の形態

度數分布は素より多種多様の形態を採るが、最大多數の位置を標準として大別すれば、(1)正常分布(又は對稱分布)、(2)非對稱分布、(3)J字型分布及び(4)双峯分布、の四つに分つ事が出來る。

(1)正常分布 (Normal Frequency Distribution)

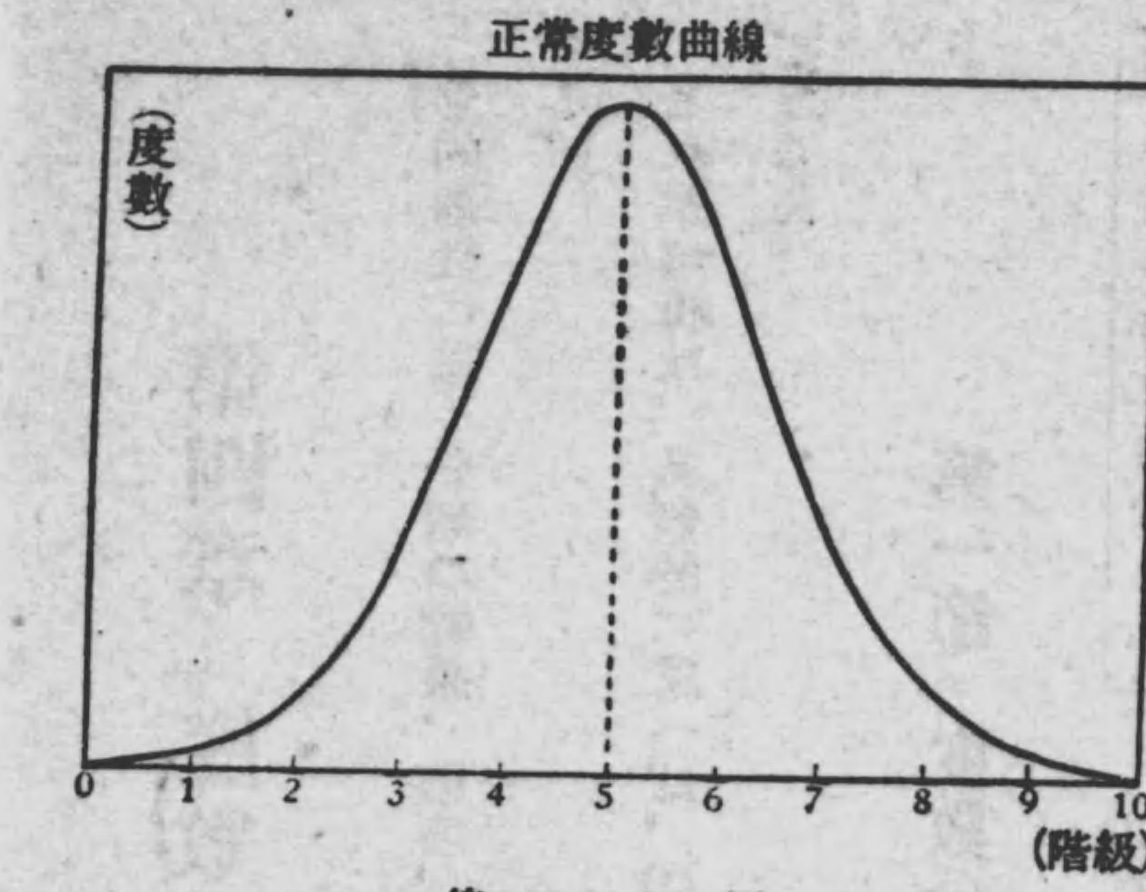
十箇の銀貨を同時に投擲し、地面に落ちたとき何枚が裏を出してゐるかを百回續けて行つた結果は上の如くであつた(各數字は各回の結果即ち現はれた裏の枚數を示す)。

6	4	6	8	6	7	3	8	6	5
4	1	6	1	6	5	7	5	7	5
5	5	6	4	4	4	6	8	4	4
5	2	6	4	3	5	5	6	5	3
4	5	4	5	5	4	6	5	3	5
5	3	5	5	3	4	6	5	5	7
4	4	3	5	4	5	4	6	8	6
5	4	2	5	8	5	5	4	5	5
5	7	6	6	7	4	5	3	3	4
5	3	5	6	2	6	3	5	5	6

これを要約すれば次の如くである。

裏の枚數	0	1	2	3	4	5	6	7	8	9	10
度　數	—	2	3	12	20	33	19	6	5	—	—

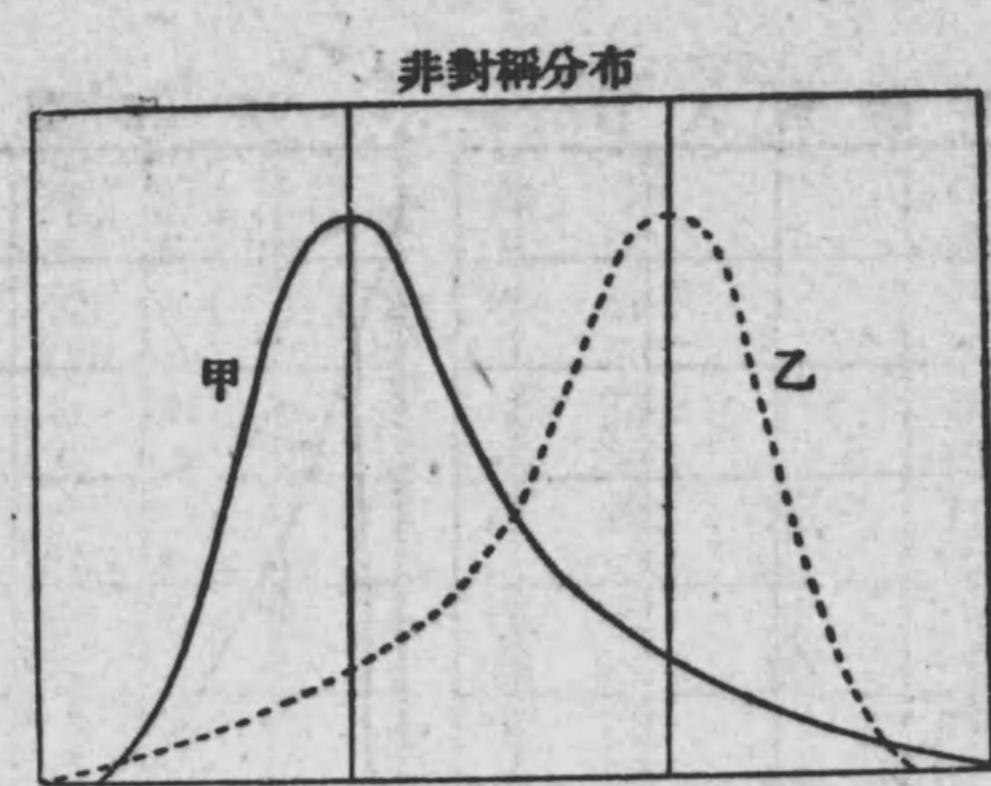
右表によれば裏が一枚も現はれなかつたとき及び九枚又は十枚が裏だつたときは一回もなく、反之、裏が五枚のときは三十三回もあり、それに次では四枚、六枚、三枚、七枚等である。これは五枚の場合を中心として左右に略〻對稱的に分布してゐるといふ事である。擲錢に於て表と裏とが現はれる回數は全く同一なのが自然であるから、裏が五枚のとき（卽ち表が五枚のとき）が最も多いのは當然である。全部が必ずしもこの結果を示さないのは、偶然の作用に基くものであるが、偶然なるものは極端なものほど少いのが確率理論上の定理であるから、五枚より四枚、四枚よりも三枚の場合がより少いのは、これ亦當然である。そして裏の四枚（卽ち表の六枚）現はれる確率は表の四枚（卽ち裏六枚）の現はれる確率と差のあるべき筈はないから、五枚を中心として左右に漸減的對稱を採る事は理論的に期待出來よう。そして投擲の回數を增加すればするほど偶然の作用は次第に相殺されるから、益々對稱的な分布を示すに至るのである。斯かる分布を正常分布といひ、それを圖示した線を正常度數曲線（Normal Frequency Curve）といふ。



第一圖

この形態の度數分布は理論的に幾多の重要な問題を提供してゐる。その說明には多分の豫備的知識を必要とするから、私は寧ろ遙か後章に於て別に論及したいと考へてゐる。實際に現はれる統計系列は大部分は非對稱的分布を示してゐる。併し身長の如き生物的現象には著しくこれに接近した分布を示すものがある。

(2) 非對稱分布（Skewed Distribution, Asymmetrical Distribution.）



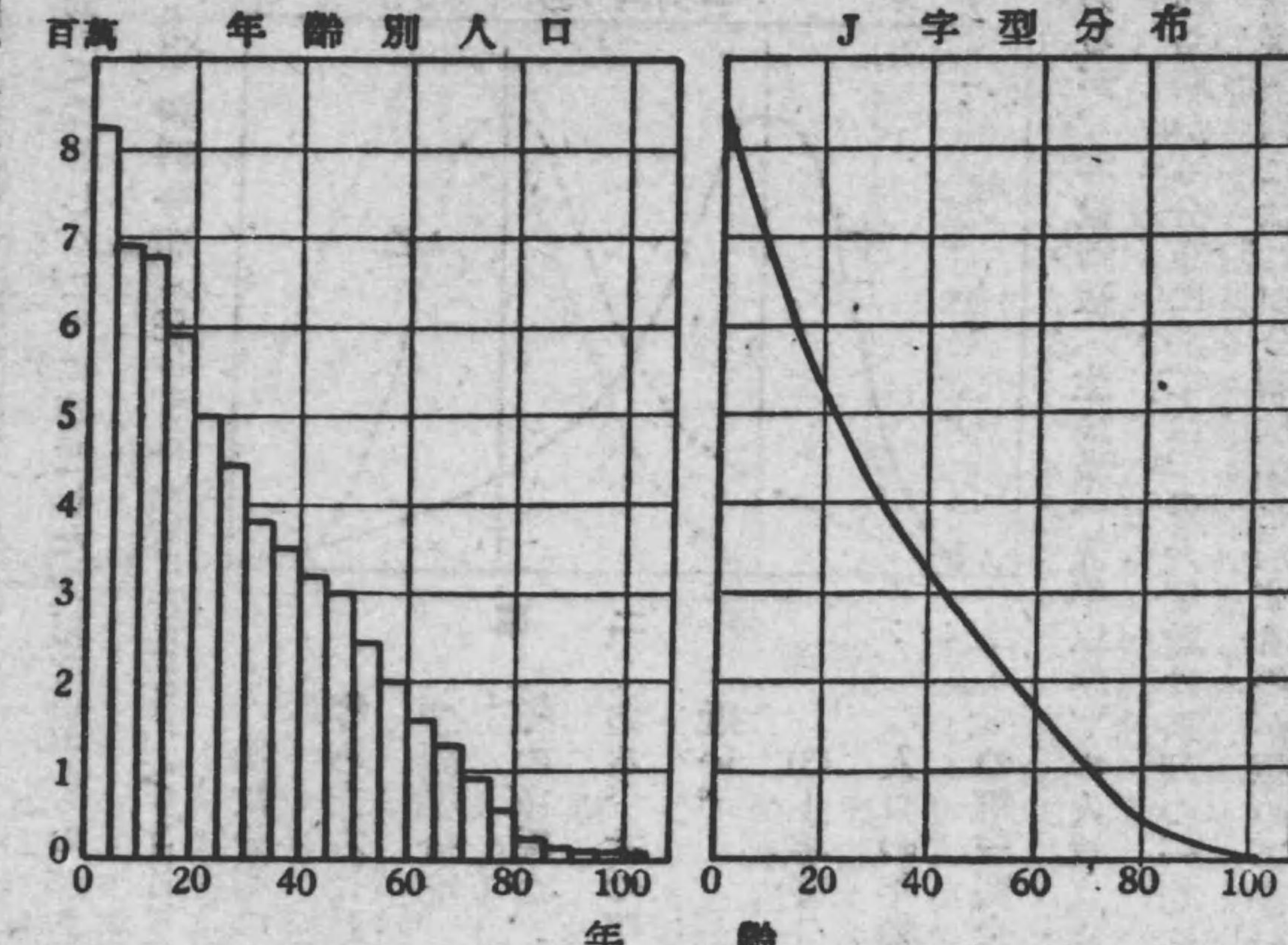
第二圖

度數分布の大部分は、最大度數を中心とした左右對稱の形態を示さないで、寧ろ左右何れかに偏倚を示すのが常である。上揭第二圖の甲は正の（又は右の）非對稱分布、乙は負の（又は左の）非對稱分布といふ。何れも最大度數卽ち曲線の頂點が、全階級卽ちX軸の中央から或ひは左に、或ひは右に偏在するのである。次の平均値の計算に繰返へし使用する賃銀別職工の系列は、甲の形態に近い。

(3) J字型分布（J Distribution）

人口の年齡別構成を見るに、若い年齡階級ほど人員の多いのが普通である。次の第三圖は本邦の大正十四年の實數をヒストグラム及び曲線で示したものであるが、零歲階級（生後滿一箇年以內）の人口（度數）が最も多く、年齡の增進と共に、度數は漸次減少してゆく。卽ち斯かる分布に於ては、最大度數は系列の一端に在り、前述の正常分布又は非對稱分布とは全く趣きを異にする。經濟系列にはこの分布狀態を示すものが極めて多い。使用職工別工場數、所得金額別第三種所得稅納稅人員數の如きはこの例である。唯だ注意すべきは、斯かる分布を示す系列でも、もし調査の範圍を擴大したり、又は階級の間隔を

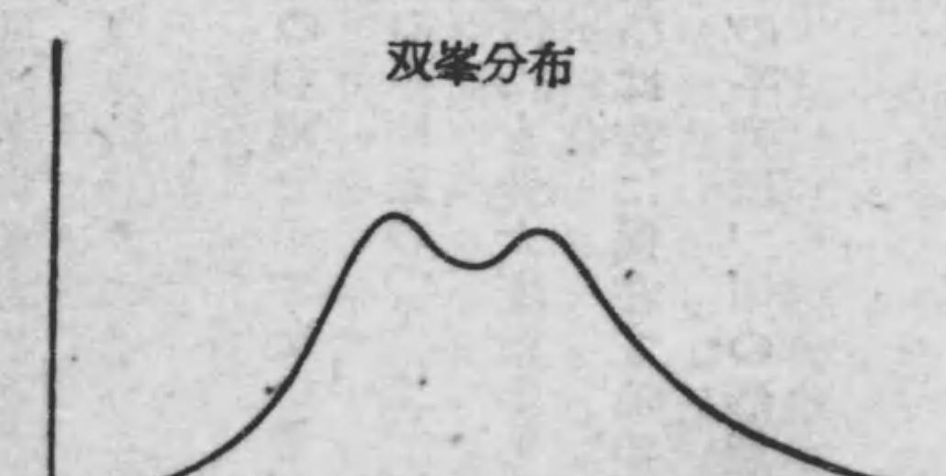
小さくすれば非對稱分布になる場合の少くない事である。



第三圖

(4)双峯分布 (Bi-modal Distribution)

いま迄述べた分布は何れも系列の一箇所を頂點として漸次下降してゆく曲線形であつたが、實際には一系列に二つ又は三つの所謂「複峯」を示すものも少くないのである。年齢別死亡數は乳兒階級と老年階級に於て大なる度數を示し、且つその間に二十歳前後の階級に又小さな突起が見られる。又人員別賃銀の如きも例へば一圓と一圓六十錢の二箇所に於て最大度數を示すが如き場合が多い。斯かる分布を双峯分布といふ。併し斯かる分布狀態は、概して異る數箇の原因の合成的結果と見られるのである。死亡に就て見るに、幼年期、青年期及び老年期の死亡原因は何れも相互に著しく相違するから、各期に特有な死亡型がある譯けで、從つてこれを一括した全人口死亡數が數箇の峯を示すのは毫も不思議ではない。同様に賃銀は男工と女工とでは全く異つた分布を示し、兩者の最大度數は男工では大なる階級に、女工では小なる階級にあるから、男女工を一括した人員別賃銀ではこの二つの階級にそれぞれ峯が現はれるのである。故にかゝる双峯分布に接した場合には、それが異る二つの系列の合成だと推察して先づ差支へがないであらう。



第四圖

第二節　統計的平均値の測定

人間社會の諸現象の裡に自然法則を探究する事を以て統計學の課題と考へたケトレーは、かゝる社會の自然法則は人間の凡ゆる心理的・道德的・生理的諸性質を平均して得られる「平均人」によつて示されると考へた（第二章一四頁）。彼に從へば箇々の人々の諸性質は各種の原因によつて相互に多少の相違を示すが、而もこれを集團として觀察すれば相互の差異は相殺されて了ふから、「平均人」は畢竟社會の本質を示すもの、換言すればそれ自身が社會の自然法則だといふのである。彼の自然法的社會觀は素より否定さる可きものであるが、併し此處に吾人は平均なる概念の眞髓に觸れる事が出來る。即ち平均は箇々の差異を相殺し事物の本質を表顯する手段なのである。平均とは事物の正常の姿を示すもの、從つて錯雜せる社會現象の研究に於て不可缺の概念と言はねばならぬ。現代統計學界の巨人ボーレー教授が「統計學は正しく平均の學であると稱してよからう」といひ、又エッジワース教授が「統計學は社會現象に就いて行はれる平均に關する學問である」といふのも强ち誇張とのみ逕ゆべきではない。平均（中數ともいふ）とは與へられた數列を一箇の數値に壓縮したもので、これが算出には幾多の數學的方法があ

る。この場合、方法が異るに從つて、得られる平均も異るのが原則である。統計的平均は勿論數學的と同樣の方法によつて求められるものであるが、併し兩者の性質は決して同一ではない。數學的平均値は純形式的なもので、その正否は全く計算の正否のみに懸るに對し、統計的平均値は——他の一切の統計的計算と同じく——數列の性質によつて規定された內容的なものでなければならぬ。これは數學の取扱ふ數量が抽象的形式的なるに對し、統計學の取扱ふ數量は具體的內容的であり、又あらねばならぬ當然の結果である。換言すれば、統計的平均の算出に於ては、單に計算の正確のみならず、同時に算式の種類の嚴選を必要とし、且つ斯くて得たる平均値は常に資料の性質によつて吟味されねばならないのである。これを忘れて統計的平均値に過度の妥當性を附與すれば、曾てケトレーの陷つた「平均人」的な自然法的社會觀に導かれざるを得ないであらう。

統計的平均は廣義に解釋すれば動態的平均、例へば時系列に於けるトレンド（次章參照）等をも包含するが、後者は意味も算式も著しく異るから、こゝでは主として度數分布の平均に限定したい。そしてこの場合の算式そのものは、一般數學的平均に於けると毫も異るものではない。一般に平均は計算的平均と位置的平均とに分れる。前者は算術平均、幾何平均、調和平均の如き、運算による平均であり、後者は並數、中位數の如き、度數分布の形態による平均である。以下この五つを順次說明しよう。

(1)算術平均（Arithmetic Mean = M_a）

與へられた數列の各項の和を項數で割つたものを算術平均といふ。5、8、10の算術平均（M_a）は $\frac{5+8+10}{3} = 7.66\cdots\cdots$ であり、一般に

$$M_a = \frac{X_1+X_2+X_3+\cdots\cdots X_n}{N} = \frac{\Sigma(X)}{N}$$

と記す（Σはシグマと讀み、總和を意味する）。然るに度數分布に於ては各階級の度數は異るから、これを考慮に入れねばならぬ。その算式は

$$M_a = \frac{\Sigma(f\times 中點)}{\Sigma(f)}$$

工職別給週

階　級	中點	度數(f)	f×中點
円4—7.99	円6	5	30
8—11.99	10	15	150
12—15.99	14	46	644
16—19.99	18	68	1224
20—23.99	22	58	1276
24—27.99	26	32	832
28—31.99	30	22	660
32—35.99	34	10	340
36—39.99	38	2	76
40—43.99	42	2	84
44—47.99	46	0	0
48—51.99	50	1	50
		261	5366

いま二六一人の職工を週給によつて分類した上表から、一人平均の週給を求めようとする。これが爲には先づ全體の週給總額を求め、これを職工數で割ればよい。然るに例へば第一の四圓乃至七圓九九錢の階級は五人の職工（卽ち度數）から成る。この五人の週給合計が幾何かは確定されない。何となればこの五人の賃銀は何れも四圓から七圓九九錢までの間である事しか判らないから。故にその合計を最も合理的に決定するには、この五人の週給が孰れも丁度階級の中央に當る六圓だと假定する外はない。これを中點（Midpoint = M. P.）といふ。斯くすればこの五人の賃銀合計は 6圓×5＝30圓 となる。同樣の計算を各階級に施して階級毎に賃銀合計を求め、次にこれら合計を合計すれば、二六一人の賃銀總計五三六六圓が得られる。この總計は實より實數とは多少の差が起り得る筈であるが、近似數たる事は認められよう。故にこれを前記公式に當嵌めれば次の如くなる。

$$M_a = \frac{\Sigma(f\times 中點)}{\Sigma(f)} = \frac{5366}{261} = 円20.56$$

$M = \frac{\Sigma fx}{\Sigma f}$

簡便法　算術平均と各項との差（偏差、Deviation）の合計は零となる。(註) これを利用して簡便計算法が行はれる。前例の5、8、10の平均を求めるに先づ假りに或る數（暗算によつてなるべく眞の平均に近い數をとるがよい）例へば8を假設平均とし、これと各項との偏差を求めれば　$5-8=-3$, $8-8=0$, $10-8=+2$　となる。若し8が眞の平均であれば、偏差の合計は零となるが、この例に於ては　$(-3)+(0)+(+2)=-1$　となつて零にはならぬ。故に偏差を零とする爲には、換言すれば眞の平均を求める爲にはこの偏差の合計を項の數で除したものを假設平均に加へればよい。即ち　$8+\frac{-1}{3}=8-0.33\cdots\cdots=7.66\cdots\cdots$　となつて、前例の結果と一致する。この眞の平均と各項の偏差は　$5-7.66=-2.66$, $8-7.66=0.33$, $10-7.66=2.33$　∴ $-2.66+0.33+2.33=0$　となる。右の原理は度數分布についても當嵌まる。即ち或る階級の中點を假設平均とし、これと各階級の中點との偏差に各度數を乗じた積の和を度數總和で割つたものを、右の假設平均に加へればよい。假設平均を十八圓として前例に適用すれば、上表の如く

階級	中點	度數 (f)	偏差 (d)	$f\times$d
¥4—7.99	¥6	5	−12	−60
8—11.99	10	15	−8	−120
12—15.99	14	46	−4	−184
16—19.99	18	68	0	0
20—23.99	22	58	4	232
24—27.99	26	32	8	256
28—31.99	30	22	12	264
32—35.99	34	10	16	160
36—39.99	38	2	20	40
40—43.99	42	2	24	48
44—47.99	46	0	28	0
48—51.99	50	1	32	32
		261	120	668

$$M_a=18+\frac{668}{261}=18+2.56=¥20.56$$

註　各項をX_1 X_2……X_n、算術平均をaとすれば、偏差の總和は　$(X_1-a)+(X_2-a)+(X_n-a)=\Sigma X-na$　然るに　$a=\frac{\Sigma X}{n}$

$$\therefore\ \Sigma X-na=\Sigma X-\frac{n\Sigma X}{n}=0$$

この事は度數分布に於ても同樣である。

欠

MISSING

いへる。併し項の數値の變化は當然平均値を變化せしむべきだとすれば、中位數は採るべきではない。要するに問題は、極端な數値が全く偶然的なものかどうかによつて決せられるのであつて、無條件にこの平均が良いとか悪いとかは云々し得ないのである。

最後に並數は、最も屢〻繰返へし現はれる數値であるから、性質上最も正常的な値と認められる。極端な數値に全く影響されぬ點は、中位數と同様である。唯だこの平均は、與へられた項數が極めて多くない限り、時には全く求められず——全項が何れも異る數値を持てば、並數は無い譯けである——たとへ求められても充分の信頼を置き難い。併し統計系列は原則上多數の項から成るものであるから、この缺陷は統計的研究に於ては根本的なものではない。

要するに以上の各平均は各〻獨自の特徴を具備し、各〻異つた數値を示すものである。その何れを使用す可きかは對象たる資料の內容的性質や項數の多寡によつて決せらる可きであり、従つて一般的にその優劣を論ず可きではないのである。その具體的問題は後章物價指數を論ずる場合に言及するであらう。

第四節 散布度

平均は系列の代表値であるが、或る平均が幾許の程度までその系列を代表してゐるかは、單に平均の種類によつて異るのみならず、同時に系列の分布狀態によつて異るのである。4、5、6の算術平均は5であり、1、5、9の算術平均も亦5である。一般に各項の數値が或る一點に集中する傾きある場合には、これを一箇の値で代表せしめる事は容易であるが、反之、集中傾向の少い場合には果して全部を代表する値があり得るかどうかは論理的に疑問であら

$\frac{4+5+6}{3}=5$

$\frac{1+5+9}{3}=5$

う。故に計算によつて求めた平均値が果してどの程度に代表値たる資格を備へてゐるかは、集中傾向の大小に比例する筈である。

系列の集中傾向の程度を、延いて平均値の信頼度を測定する尺度を散布度（Dispersion）といふ。即ち平均値は常にその散布度によつて意味が確定するのであつて、若しこれを怠つて單に平均のみを記せば、恰も某君の試驗は五番であると言ふだけで受驗人員を言はぬと同じ事になる。千人中の五番と、十人中の五番では、まるで意味の違ふ事は言ふ迄もなからう。單に平均は5であるといふのもこれと同じである。

斯くて度數分布の解析に於て、散布度の測定は不可缺の要件たるものであり、幾多の方法が案出されてゐる。此處ではレンジ、平均偏差、標準偏差及び四分一偏差、の四つにつき説明しよう。このうち最も合理的なものは標準偏差である。

(1)レンジ（又は極差、Range）

與へられた系列の上下兩端に位する二項の幅をいふ。4、5、6の平均が5なりといふ時、この5は4—6の幅を平均した値であり、1、5、9の平均が同じく5なりといふ時、この5は1—9の幅を平均した値である。狹い幅は集中傾向を示すから、レンジの小さい平均値としての價値が大なるを意味する。株式取引所に於ける株價の變動を發表する際、例へば昨日の新東株の高値は一三九圓〇、安値は一三八圓三なぞと言ふのはレンジに外ならぬ。レンジは方法極めて簡單であるが、單に兩端の二項の數値のみによつて決定されるから、何等か偶然の理由によつてこの兩端の項の値が法外に動いたとすれば、他の諸項の値は動かずとも、レンジは著しく變化する。百の項から成る系列でも、レンジは僅かに兩端の二項のみに左右され、殘りの九十八項の値は無視される。分布狀態は斯かる單純な方法によつては到底充分に知り得るものではない。

(2)平均偏差（Mean Deviation）

算術平均又は中位數と各項との偏差の絶對値（正負の記號を無視せるもの）の算術平均を平均偏差といふ。レンジの缺陷は全部の項を考慮に入れぬ事であつた。平均偏差はこの缺陷を是正するものである。いま各項をX_1 X_2……X_n、平均（M_a又はM_e）をM、項數をN、とすれば

$$\text{平均偏差 (M.D)}=\frac{(X_1-M)+(X_2-M)+\cdots\cdots(X_n-M)}{N}$$

$$=\frac{\Sigma(d)}{N}$$

となる（dは偏差の記號である）。4、5、6の系列の平均偏差は、$\frac{(4-5)+(5-5)+(6-5)}{3}=\frac{2}{3}=0.666$であり、1、5、9の平均偏差は$\frac{(1-5)+(5-5)+(9-5)}{3}=\frac{8}{3}=2.666$となる。即ち平均偏差の小なるほど集中傾向の大なること、或ひは平均値の値値の大なることが判らう。度數表から算出する式は次の如くである。（f_1 f_2……f_nは各階級の度數）。

$$\text{平均偏差}=\frac{f_1(X_1-M)+f_2(X_2-M)+\cdots\cdots f_n(X_n-M)}{f_1+f_2+\cdots\cdots f_n}$$

$$=\frac{\Sigma(fd)}{\Sigma f}$$

$f-20.56=14.56$

$\Sigma f(x-M)f$

階級	中點	偏差	度數	度數×偏差
≒4—7.99	≒6	14.56	5	72.80
8—11.99	10	10.56	15	158.40
12—15.99	14	6.56	46	301.76
16—19.99	18	2.56	68	174.08
20—23.99	22	1.44	58	53.82
24—27.99	26	5.44	32	142.08
28—31.99	30	9.44	22	207.68
32—35.99	34	13.44	10	134.40
36—39.99	38	17.44	2	34.88
40—43.99	42	21.44	2	42.88
44—47.99	46	25.44	0	0
48—51.99	50	29.44	1	29.44
			261	1352.22

$$\therefore\ \text{平均偏差}=\frac{1352.22}{261}\fallingdotseq 5.18$$

これを從前の賃銀別職工表に適用して見よう。その算術平均は二十圓五十六錢であつた。これと各項との偏差を求むるには、當然各階級の中點をとらねばならぬ。依て計算は前表の如くなる。即ち平均二十圓五十六錢、平均偏差五圓十八錢となる。

(3) 標準偏差（Standard Deviation＝σ）

平均偏差に於ては偏差の記號を無視するの不合理を犯した。單に絶對的大さを問題とする以上、かゝる方法も許容さるべきであるが、併し數學的正確を期するならば、何等かの方法によつて是正するのが望ましい。偏差には正負の別が現はれるから、これを平方すれば合理的に總てを正の數として取扱ふ事が出來る。標準偏差とはこの理に基き、算術平均と各項との偏差の自乘の總和を算術平均したものゝ平方根を指すのであつて、一般に希臘文字σ（シグマ、Σの小文字）を用ひる。即ち各偏差をdとすれば

$$\sigma=\sqrt{\frac{\Sigma(d^2)}{N}}$$

4、5、6のσは $\sqrt{\frac{(4-5)^2+(5-5)^2+(6-5)^2}{3}}=\sqrt{\frac{1}{.3}}=\sqrt{0.333}=0.58$ となり、1、5、9のそれは同樣にして $\sqrt{\frac{32}{3}}=\sqrt{10.66}=3.26$ となる。

度數表からは、同じく度數を考慮して

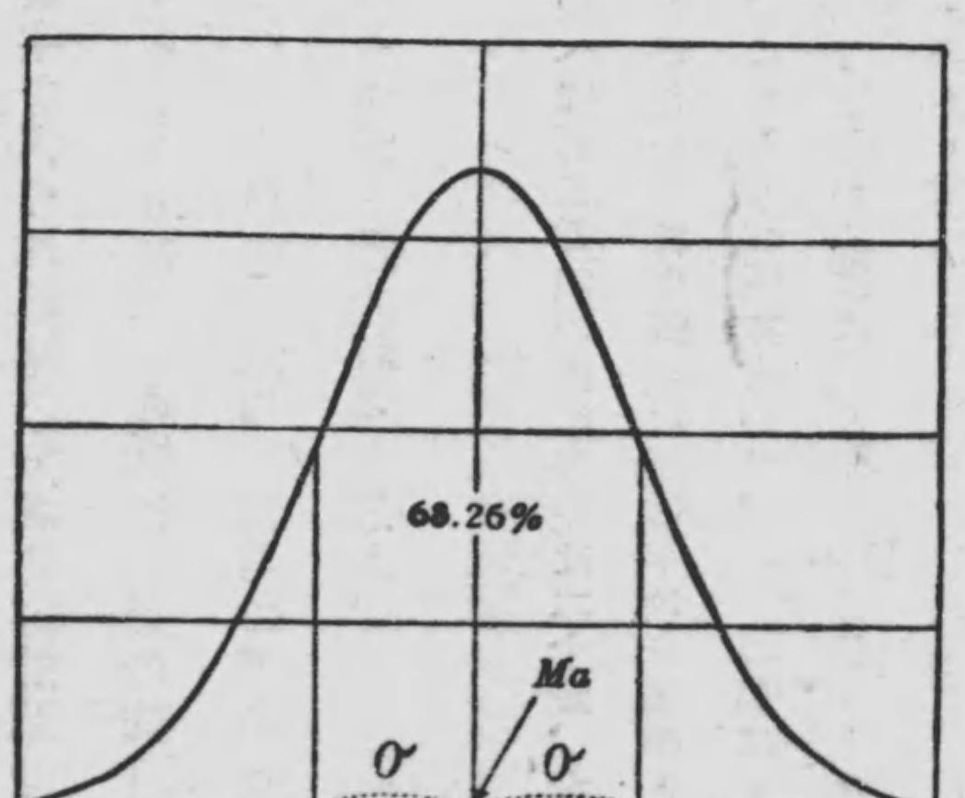
$$\sigma=\sqrt{\frac{\Sigma(fd^2)}{\Sigma(f)}}$$

これを前例に適用すれば次表の如くなる（階級の數字は省略した）。

中點	d	d^2	f	fd^2
¥6	−14.56	211.99	5	1059.95
10	−10.56	111.51	15	1672.65
14	− 6.56	43.03	46	1979.38
18	− 2.56	6.55	68	445.40
22	＋ 1.44	2.07	58	120.06
26	＋ 5.44	29.59	32	946.88
30	＋ 9.44	89.11	22	1960.42
34	＋13.44	180.63	10	1806.30
38	＋17.44	304.15	2	608.30
42	＋21.44	459.67	2	919.34
46	＋25.44	647.19	0	0
50	＋29.44	866.71	1	866.71
			261	¥12385.39

$$\therefore\ \sigma=\sqrt{\frac{12385.39}{261}}=\sqrt{47.27}=6圓87$$

標準偏差は系列内の集中傾向並びに算術平均の信頼度を示すもので、その小なるほど理想的なことは平均偏差に於けると同樣である。併しこれ以外に標準偏差によつて、分布狀態を更に確定する事が出來る。即ち正常分布（對稱的分布）に於ては、算術平均を中心としてその上下に各〻標準偏差の大さを取れば、その範圍内、即ち $M_a \pm \sigma$ の中に、全項の六八・二六％が包含され、標準偏差の二倍を取れば、その範圍内、即ち $M_a \pm 2\sigma$ の中に全項の九五・四六％が包含されるのである。假りに千人の學生の身長が正常分布なりとし（身長の如きは概して斯かる分布を示す）、算術平均が五尺三寸、標準偏差が一寸五分なりとすれば、約六八三人は 5尺3±0尺15 即ち五尺一寸五分から五尺四寸五分までの身長を有し、又約九五五人は 5尺3±2×0尺15 即ち五尺から五尺六寸までの身長を有するのである（第六圖參照）。一般の度數分布は素より理論的正常分布とは多かれ少かれ相違するから、右の數



第六圖

字がそのまゝ當嵌まる筈はないが、而も多くの場合、可成りこれに接近した數字を得られるものである。前例の賃銀別職工表に就て見るに、分布は可成り不均分であるが、而も $M_a \pm \sigma$ 即ち 20圓56±6圓87 の間に全員の約八割が含まれる。該表の十四圓乃至二十六圓の階級の度數即ち人員は二〇四人で、これは全員の七八%に當るのである。即ち平均と標準偏差とが與へられゝば、これのみから度數分布の概況を推察し得る事が了解されると思ふ。

註　變化係數 (Coefficient of Variation)

甲乙二系列の散布狀態は、直接に兩者の標準偏差を比較しても判るものではない。例へば男工賃銀の標準偏差は二十錢、女工賃銀のそれは十錢なりとしても、直接二十錢對十錢の比がその各〻の散布狀態を反映し得ない。標準偏差は算術平均を基準として算出されたものであるから、基準たる算術平均に對する比率に換算する必要がある。これを變化係數といひ、$\frac{\sigma}{M_a}$ の形によつて示される。上例に於て平均賃銀を八十錢及び五十錢とすれば、各〻の變化係數は $\frac{5}{80} \times 100 = 6.25\%$ 及び $\frac{4}{50} \times 100 = 8\%$ となり、標準偏差は絕對的には女工賃銀の方が小であつても、比較的には大なる事が判らう。

(4) 四分位偏差 (Quartile Deviation＝Q. D)

全項を四等分する各限界は第一四分位 (First Quartile)、第二四分位 (Second Quartile) 及び第三四分位 (Third Quartile) である。第二四分位は全項の中央に位するから中位數に外ならぬ。この第三四分位 (Q_3) と第一四分位 (Q_1) との差を二等分したものを四分位偏差といふ。即ち

$$Q.D = \frac{Q_3 - Q_1}{2}$$

第一及び第三四分位は、中位數を求める方法と同樣にして求められる。唯だ中央の代りに四分一及び四分三の個所に當る項を求めればよいのである。再び中位數決定の際の賃銀別職工累積度數表を用ひて計算して見よう。該表は次の

階級	度數	累積度數
¥4—7.99	5	5
8—11.99	15	20
12—15.99	46	66
16—19.99	68	134
20—23.99	58	192
24—27.99	32	224
28—31.99	22	246
32—35.99	10	256
36—39.99	2	258
40—43.99	2	260
44—47.99	0	260
48—51.99	1	261
	261	

如くであつた。中央項は百三十一人目であつた。故に第一四分位 Q_1 は六十五人目、第三四分位 Q_3 は百九十六人目である。その値は

$$Q_1 = 12圓 + \frac{46}{46} \times 4圓 = 16圓$$

$$Q_3 = 24圓 + \frac{4}{32} \times 4圓 = 24圓5$$

となるから、この二つを前式に代入すれば

$$Q.D = \frac{Q_3 - Q_1}{2} = \frac{24.5 - 16.0}{2} = 4圓25$$

四分位偏差が散布度測定の手段たり得る理由は、度數分布の散布度が增大するに從つて、四分位の距離は增大するからである。四分位偏差は第三及び第一四分位の差の半分であるから、分布が全く對稱的なときには中位數と一致するが、斯かる正常分布は理論的にしかあり得ないから、實際には多少の偏倚を示すのが常である。

第五節　非對稱度

度數分布の解析に於て、集中傾向の程度卽ち散布度を測定する必要あると同じく、分布の對稱の程度を測定する必要もある。全く對稱的な分布は理論的にのみ可能であつて、實際には多かれ少かれ非對稱的分布を示すからである。

非對稱度は歪度 S_k とも呼ばれ、その測定には二つの方法がある。

(一) 完全に對稱的な分布に於ては算術平均、中位數及び並數は全く一致する。非對稱的分布に於てはこの三者は相互に異り、且つ非對稱度が增大するに從つてその差は益〻增大する。算術平均は極端な數値によつて最も強く影響

されるから、非對稱度の増大と共に、算術平均は益々並數又は中位數（兩者は極端な數値の影響を全く受けぬ）から遠ざかる。故に並數又は中位數と算術平均との距離の大小は非對稱度の尺度たる事明かである。

然るに非對稱度は元來相互比較の爲に必要なのであるから、右の方法を以てしては次の二つの困難に遭遇する。即ち、(1)比較さるゝ甲乙二系列の單位は必ずしも等しくない。(2)散布が大なる系列に於ては如上の距離は當然增大する。故に甲乙の非對稱度を比較する爲には、これを同一の基準に換算せねばならない。これには如上の距離を、それぞれの標準偏差に對する比率に書改めればよいのである。即ち

$$S_k = \frac{m_a - M_o}{\sigma}$$

或ひは、既に説明した通り、略々 $m_o = m_a - 3(m_a - m_e)$ であるから、これを右式に代入すれば

$$S_k = \frac{3(M_a - M_e)}{\sigma}$$

賃銀別職工表に就て見れば、m_a ＝20圓56, m_e ＝19圓82　標準偏差は 6圓87 であつたから

$$S_k = \frac{3(20.56 - 19.82)}{6.87} = 0.32$$

この〇・三二は比率であつて、三十二錢の意味でない事に注意せられたい。上記二つの公式のうち、素より第一式、即ち並數を用ふるのがより適當である。これをピアソニアン非對稱係數（Pearsonian Coefficient of Asymmetry）といふ場合もある。

（二）非對稱度は四分位の位置からも測定される。對稱的分布に於ては第一及び第三四分位は中位數から等距離に在るが、非對稱分布に於ては然るを得ないから、この距離の大小によつて非對稱度を測定し得る理である。この距離を、四分位偏差に對する比率として示せばよいのであつて、即ち

$$S_k = \frac{(Q_3 - M_e) - (Q_1 - M_e)}{Q.\ D}$$

前例に適用すれば

$$S_k = \frac{(24.24 - 19.82) - (16.00 - 19.82)}{4.25} = 1.93$$

この二つの方法の何れが勝れたりやは困難な問題である。第一の並數による方法は非對稱が小なる場合にも明瞭に指示すると言はれてゐるが、一般に並數の決定は困難な場合の少くない事を考へれば、中位數による代用法か又は四分位法を用ふる必要があらう。

第五章　時系列の解析(一)

第一節　時系列に現はれた運動形態

時の經過と共に變化する現象を、時の順序に從つて數字的に把捉したものを、時系列（Time-Series）といふ。時系列に現はれた變動は、大部分の場合、決して單一な運動ではなく、數箇の異種的運動の綜合的結果なのである。統計學の最も重要な課題の一つは、與へられた時系列の示す變動を其の數箇の構成的運動に分析する事であつて、これを時系列の解析といふ。

時系列に於ける變動を惹起せしめる運動は、形式的には次の四つに分類される。

1、長期間に亙る一般的方向、即ち長期傾向（發展傾向又は趨勢）

2、一年以內に一巡する季節變動

3、景氣の上昇及び下降の交代的運動、即ち循環運動（景氣變動）

4、不規則變動（偶然的變動）

この分類はハーヴァード景氣研究所の採つたもので、該研究所が時系列の統計的解析に於て最も秀拔な業績を擧げた結果、その後の統計的研究は大部分この分類を踏襲してゐる。併し該研究所に對立するベルリン景氣研究所に於て

は、右分類を餘りに形式的なものとして非難し、その所長ワーゲマン博士は次の如く言つてゐる。「ハーヴァードの分類は系列を主として形式的に見るといふ點から爲されたのであるが、それはその背後に横はる整序原理が不明瞭なるが故に滿足出來ない。吾々が……構造變化と景氣變動とを對立せしめるといふ立場からすれば、寧ろ運動形態はこれを次の如く分類するを以て可とする。

1、一回限りの變化

(a)非連續的（發展挫折）——狀態變化、崩壞、分裂

(b)連續的（發展）——生長、變革、後退

2、週期的變動

(a)一定のリズムをもつもの（季節變動）

(b)一定のリズムをもたざるもの（狹義の景氣變動）」

斯く形式的と實質的の二箇の分類があるが、併し實際上は兩者は全く同一視してよいのである。蓋し後者の1の(a)は不規則變動、(b)は長期傾向、2の(a)は季節變動、(b)は循環運動に外ならぬからである。そして事實ベルリン景氣研究所の採用する統計的解析法は、ハーヴァードのそれと別に異るところはない。故に分類に關する理論的論爭を別とし、問題を單に統計的解析法に限定する限り、寧ろ一般に普及せるハーヴァード式分類に據る方が便利であらう。以下右四形態の運動を順次に分析して見よう。尤もそのうちの不規則變動を測定する方法に就ては、適當なものが殆どないので、深く立入る事は出來ない。

第二節　長期傾向の測定

長期傾向運動（Secular Trend）なる概念は元來天文學に發せるもので、何萬何億年といふ極めて悠久な期間に亙つて起る運動形態、例へば遊星軌道の變化の如きものを意味する。然るにこれが社會現象の研究に轉用された場合には遙かに短い期間に限定されるに至ることは止むを得ない。後者に於ては、單に比較的長期に亙る、或ひは循環運動よりも長期に亙る、發展運動を意味するのみで、可成り漠然とした概念である。經濟系列の解析に於ては普通に循環運動は三、四年乃至、七、八年の周期を持つものとされてゐるから、それより長い、例へば十年、二十年乃至それ以上に亙る一般的運動形態は長期傾向運動と認められる。併し後に述べるが如く、循環運動の特殊形態として所謂「長期波動」なる五十年前後に及ぶ大循環運動が幾多の經濟系列について發見されるに至つてから、兩者の限界は益々曖昧となつて來た。

併しいま一經濟系列、例へば明治初年以降の本邦輸出貿易額連年統計を見れば、景氣の良否、新市場の獲得、舊市場の喪失、生産技術や配給組織の變化等によつて極めて不規則な數字を示してはゐるが、併しこの間、大體に於て時と共に發展向上し來つた事は爭へない。

此處に「大體に於て」といふのは、畢竟「平均的に」といふ意味である。そして平均とは、既に説明した通り、事物の正常的狀態を指すに外ならないから、從つて右貿易額の正常的發展狀態は上昇的だと言つてよい事になる。そして年々の計數は、その時々の事情によつてこの正常狀態の或ひは上に、或ひは下に在るわけで、換言すれば右の正常

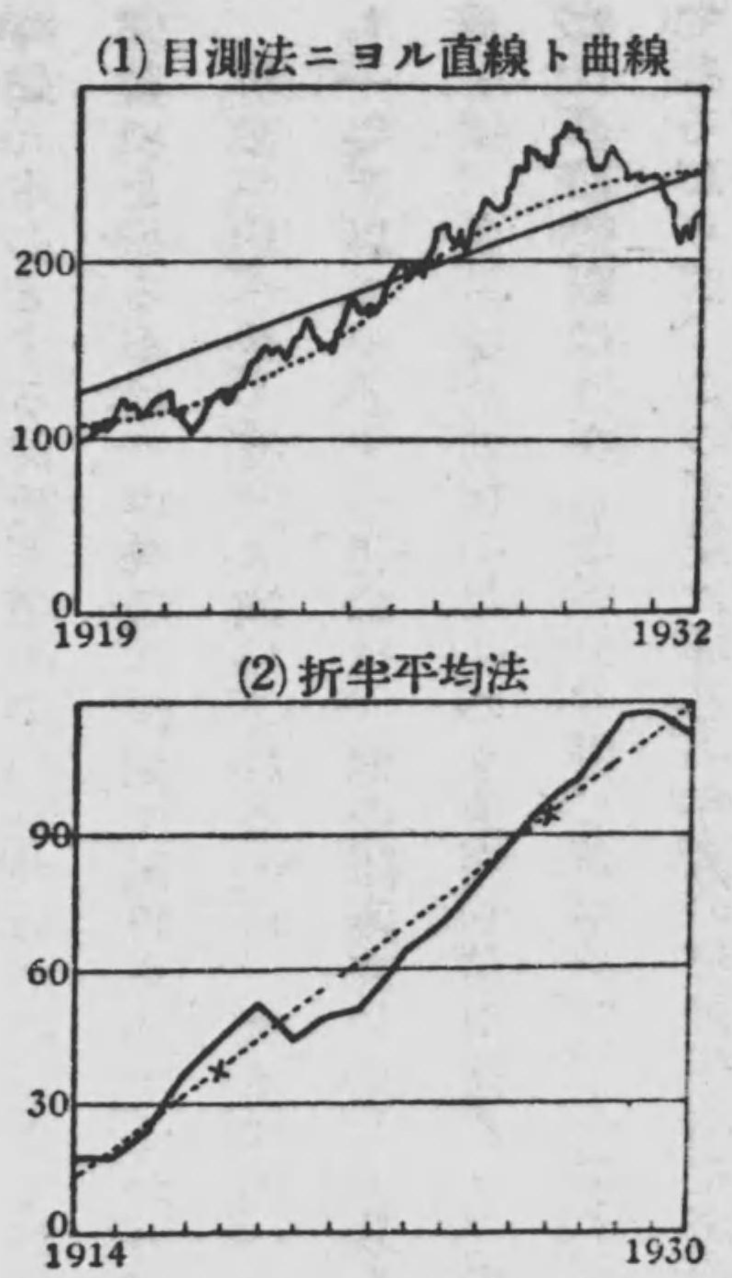
狀態はその時々の計數を判定する基準となるものである。唯だ時系列に於ける斯かる平均値は、度數分布の場合のそれと異り、一箇の代表値ではなくて一箇の代表線である。これは時の經過を取扱ふ時系列の性質から容易に窺へるであらう。この代表線を傾向線といふ。

時系列解析に於ける傾向線の算出は二箇の目的を有する。一は上に述べた通り、一系列の長期に亙る發展傾向を知る事であり、他の一つは發展傾向をば他の變動形態卽ち季節變動及び不規則變動と共に原系列から除去し、以て循環運動形態卽ち景氣變動樣態を求める事である。卽ち前者は傾向線そのものを目的とし、後者は別の目的に對する手段たる事を目的とするものである。併し目的は斯く異つても、傾向線算出の方法そのものは原理的には毫も異るものではない。尤も後者の目的の爲には可成り便宜的方法の用ひられる事もあるが。……(この問題は、「經濟統計」に於て取扱ふであらう。)

傾向線算出に關する統計技術は極めて複雜多岐に亙り、本書に於ては餘り深く觸れる事は出來ない。此處では最も普通に用ひられる若干の方法を例示的に說明しよう。

(一) 目測法 (Free-hand Method)　最も簡單に傾向線を求めるには、與へられた時系列のグラフを一覽し、上昇又は下降の一般的方向、卽ち平均線と認められる直線又は曲線を目測によつて描けばよい(第七圖參照)。目測は當事者自身の判斷に左右され、客觀的標準を缺くから、殆ど總ての統計學教科書に於て、この方法を以て最も不完全なものとして取扱つてゐる。

併し與へられたグラフに餘り不規則な凹凸がない場合には、主觀の働く範圍は狹くなり、誰が描いても略〻同じ線



第　七　圖

が得られるから、必ずしも後段に述べるが如き複雜な數學的方法を用ひる必要はなからう。殊に熟練者が描けば、數學的方法による所謂「合理的傾向線」と極めて類似した傾向線が求められるものである。一般に主觀的判斷なるものは、後に說明する通り、最も機械的な數學的方法に於てさへも、著しい程度に介入し來るのであつて、決して目測法のみの持つ缺陷とは言へないのである。

(二) 折半平均法 (Semi-Average Method)　目測法と共に最も簡單な方法として折半平均法なるものがある。與へられた全期間を二等分し、先づ前半期の各項の平均を求め、これを前半期中央時點の坐標點とする。同樣にして後半期中央時點の坐標點を求め、この二箇の坐標點を通る直線を引くのである。この方法は甚だ機械的で、主觀的判斷の働く餘地は全くない。この點に於て目測法とは正反對の地位に立つものといへる。この方法は如上の二箇の坐標點が前後各半期のそれぞれ適當な代表値たる場合には、これによつて略〻信賴するに足る傾向線を求める事が出來る。併し平均の本質上、若し與へられた原系列中に何等か異常な値の項が介在する場合には、これによつて平均値が著しく歪められ、從つて適當な代表値たるの資格を失ふに至る。卽ちこの方法は、不規則變動の多分に含まれてゐる系列については到底適用し得られないのである。

常値の謂ひなる事を明かにした。故にいま或る年を中心とする數年間の平均を求めれば、この平均値をその年の正常値と認めて差支へなからう。然らば次にはその翌年を中心とする數年間の平均によつて、その次には翌々年を中心とする數年間の平均によつて、それぞれの年の正常値が決定される筈である。いま問題とする傾向値とは長期間內の各時點の正常値に外ならぬから、上記の如き方法によつて求めた各年の正常値は傾向値と同じ性質のものであり、從つてこれら正常値を順次に連結したものは傾向線と認められるわけである。移動平均法はこの理由に立脚する。

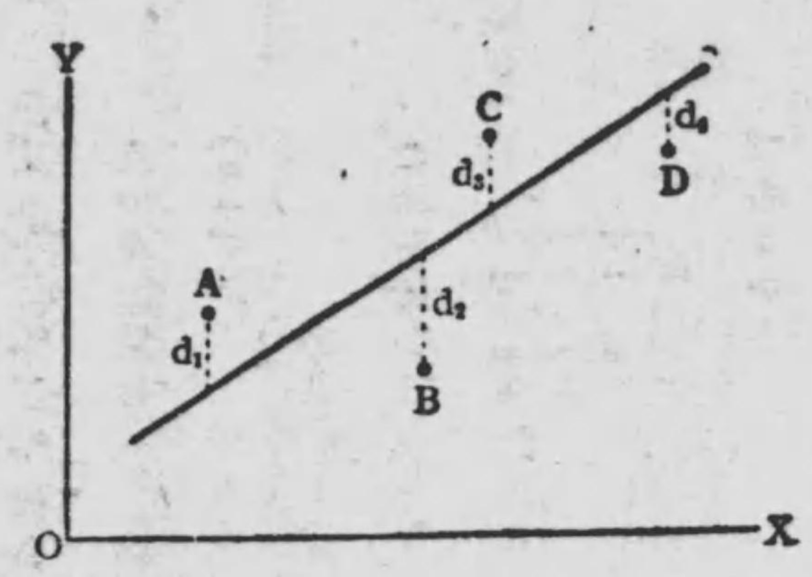
（三）移動平均法（Moving Average Method）　「平均」を論じたときに、統計的平均値とは代表值或ひは正

與へられた年別統計の或る項をU_n、その前後の隣接項を各々U_{n-1}及びU_{n+1}とすれば、その算術平均$\frac{1}{3}(U_{n-1}+U_n+U_{n+1})$は中央項$U_n$の正常値である。次に一項をづらして$\frac{1}{3}(U_n+U_{n+1}+U_{n+2})$を求めれば、右と同じ理由でこの平均値は次の$U_{n+1}$項の正常値とならう。同様の計算を最終項まで續ければ「三年移動平均」を行つた事になる。この例では或る時點を中心とする前後三項の平均を求めたのであるが、五項でも七項でも理由は同じである。（四項とか六項とかの偶數項を採る事は、中央項の決定に困難を感ずるから、出來るだけ奇數項を採るのがよい。）

移動平均が最も有效に適用されるのは、次章に述べる季節變動測定の場合である。凡ゆる變動の中で季節變動、即ち一年を以て一巡する運動は、他の諸運動に比して最も規則的である。故に月別統計について十二箇月移動平均法を施せば、各月毎に前後一年間の季節變動が相殺され、從つて季節變動の除去された値が得られるのである。年別統計には季節變動は介在しないが、併し數年を以て一巡する循環運動は介在する。故に若し循環の期間（週期）が一定してゐるならば、この週期に等しい年數の移動平均を施せば、右と同様の理由によつて、この循環運動の除去された値が得られるであらう。(註)例へば週期が正確に七年とすれば、七年移動平均を行へばよい。併し實際には斯かる正規的な循環運動は、少くとも經濟的現象に關しては先づ無いと言はねばならぬ。斯かる場合には移動平均法によつて循環運動を除去する事は不可能となり、從つてこの方法は實際の見地からは適當な方法ではないのである。

註　實は單に週期のみならず、振動の幅も亦等しい事を必要とする。

（四）曲線の當嵌め（Curve-fitting）　xに關する一次式　$y=a+bx$　は直線、二次式　$y=a+bx+cx^2$　は抛物線なる事などは、圖示法の章に於て說明した。然らば直線的傾向線は　$y=a+bx$、抛物線的傾向線は　$y=a+bx+cx^2$　の方程式を以て示される筈であつて、これを求める手續、即ち平面上に與へられた數箇の座標點を縫ふ最も近似的な直線又は曲線の方程式を求める方法を一般に「曲線の當嵌め」といふ。



第八圖

時系列に直線的傾向線　$y=a+bx$　が引かれたとすればyは縱座標即ち每年の計數（生產額、物價、職工數等）を、xは橫座標即ち年の經過を示すもので、從つてこの兩者は原系列に於て與へられてゐる數字である。故に問題は右式に於ける a 及び b の大さを決定する事に歸着する。これが手續は最小自乘法（Method of Least Squares）なる特殊の數學によつて得られた$\begin{cases}\Sigma Y=Na+b\Sigma X\\ \Sigma XY=a\Sigma X+b\Sigma X^2\end{cases}$なる公式を用ひればよいのである。(註)

註　上圖の如く平面上に町村A、B、C、Dが散在せるとき、これらに最も近く直線道路を設けんとする場合、最小自乘法に從へば、各町村からの距離　$d_1\,d_2\,d_3\,d_4$　の自乘の總和

$d_1^2+d_2^2+d_3^2+d_4^2$ 即ち $\Sigma(d^2)$ を最小ならしめるが如き條件を具へた道路が求める道路である。

Aの座標點を $(x_1 y_1)$、Bのそれを $(x_2 y_2)$……等とすれば、直線 $y=a+bx$ までの距離 $d_1 d_2$……等は $y_1-(a+bx_1)$, $y_2-(a+bx_2)$……である。故に自乘の總和をSとすれば

$$S=[y_1-(a+bx_1)]^2+[y_2-(a+bx_2)]^2+\cdots\cdots$$

これを最小ならしめるには、微分學の原理に從つて、微係數 $\frac{\delta s}{\delta a}$, $\frac{\delta s}{\delta b}$ を零ならしめればよい。然るに

$$\begin{cases}\frac{1}{2}\cdot\frac{\delta s}{\delta a}=\Sigma[Y-(a+bX)]=\Sigma Y-Na-b\Sigma X\\ \frac{1}{2}\cdot\frac{\delta s}{\delta b}=\Sigma[Y-(a+bX)]X=\Sigma XY-a\Sigma X-b\Sigma X^2\end{cases}$$

であるから

$$\begin{cases}\Sigma Y=Na+b\Sigma X\\ \Sigma XY=a\Sigma X+b\Sigma X^2\end{cases}$$

となる。

この聯立方程式からa及びbを求めて、その値を $y=a+bx$ に代入すれば、求める直線が得られるのである。

この公式のXは年度を、Yは各年の計算を、Nは項の數を示す。直ちに次の實例について運算を行つて見よう。これは一九一六年から一九三〇年に至る十五年間の米國に於けるアルミニューム生産額（單位百萬ポンド）を示したもので、從つて N=15 である。Yは各年の生産額、Xは一九一六年……一九三〇年を示すが、Xは便宜上一九一六年を零、一九一七年を一、……一九三〇年を一四と書換へる。（年は一年毎に一つ宛殖えて行くから、一九一六、一九一七といふやうな原數字をそのまゝ用ひずとも、上の如くして容易に年次を判斷し得るからである。例へば2の年が

年次	X	アルミニューム生産額（單位百萬封度） Y	XY	Y^2
1916	0	110.2	0	0
1917	1	143.3	143.3	1
1918	2	143.3	286.6	4
1919	3	134.5	403.5	9
1920	4	138.0	552.0	16
1921	5	55.0	275.0	25
1922	6	74.0	444.0	36
1923	7	129.0	903.0	49
1924	8	150.0	1200.0	64
1925	9	140.0	1260.0	81
1926	10	145.0	1450.0	100
1927	11	260.0	1760.0	121
1928	12	210.0	2520.0	144
1929	13	225.0	2925.0	169
1930	14	229.0	3206.0	196
	$\Sigma(X)$ =105	$\Sigma(Y)$ =2186.3	$\Sigma(XY)$ =17328.4	$\Sigma(X^2)$ =1015

一九一八年を意味する事は誰にも明かであらう。さて右公式に代入すべきものは年次の數(N)、年次の總和(ΣX)、各年生産額の總和(ΣY)、年次と該年次の生産額の乘積の總和(ΣXY)及び年次の平方の總和(ΣX^2)である。即ちこれらを表の上で計算すれば、上表の最後の行の各數字が得られる。

これを公式に代入すれば

$$\begin{cases}2186.3=15a+105b\\ 17328.4=105a+1015b\end{cases}$$

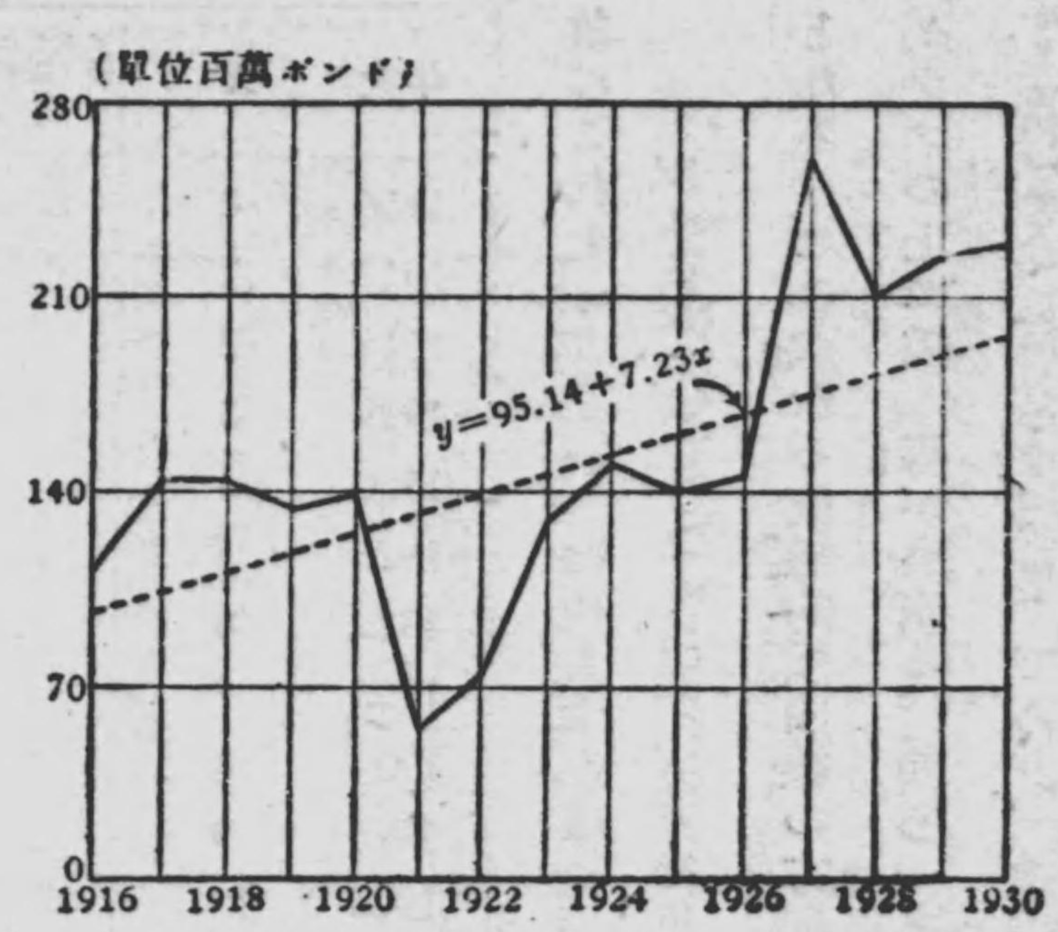
となる。これを解けば

$$\begin{cases}a=95.14\\ b=7.23\end{cases}$$

故にこれを $y=a+bx$ に代入すれば

$$y=95.14+7.23\,x\cdots\cdots\cdots\cdots\cdots\cdots(1)$$

となる。これ求むる直線的傾向線の方程式である。方程式が得られゝば、この x に年次を代入する事によつて實際にグラフの上に描く事が出來る。即ち一九一六年の y の大さ（即ち傾向線の高さ）は x に零を代入すれば得られるから、$y=95.14\times7.23\times0=95.14$ であり、次の一九一七年のそれは $y=95.14+7.23\times1=102.37$ である。直線を描くには、



第　九　圖

二點が決定されゝば、これを連結すればよいから、この二點即ち 95.14 と 102.37 を連結し、更にこれを延長しさへすればよい。第九圖の點線は即ちこれを示したものである。

併し本例の如く、年數（N）が奇數なるときは、最初の一九一六年を零とする代りに、期間の中央即ち一九二三年を零とすれば、計算が甚だ容易となる。一九二三年を零とすれば一九二二年は-1、一九二四年は$+1$となり、同様にして一九二一年は-2、一九二五年は$+2$……一九一六年は-7、一九三〇年は$+7$となり、從つて總和 $\Sigma(X)$ は $(-7)+(-6)+\cdots\cdots(-1)+0+1+\cdots\cdots6+7=0$ となる。即ち公式の $\Sigma(X)$が總て零となるから、次の如き簡單なものになつて了ふ。

$$\begin{cases}\Sigma(Y)=Na\\ \Sigma(XY)=b\Sigma(X^2)\end{cases}$$

前例についてこれを適用すれば、次頁の表の如くである。

年次	X	Y	XY	X^2
1916	−7	110.2	−771.4	49
1917	−6	143.3	−859.8	36
1918	−5	143.3	−716.5	25
1919	−4	134.5	−538.0	16
1920	−3	138.0	−414.0	9
1921	−2	55.0	−110.0	4
1922	−1	74.0	−74.0	1
1923	0	129.0	0	0
1924	1	150.0	150.0	1
1925	2	140.0	280.0	4
1926	3	145.0	435.0	9
1927	4	260.0	640.0	16
1928	5	210.0	1050.0	25
1929	6	225.0	1350.0	36
1930	7	229.0	1603.0	49
	$\Sigma(X)=0$	$\Sigma(Y)=2186.3$	$\Sigma(XY)=2024.3$	$\Sigma(X^2)=280$

これらの數字を代入すれば

$$\begin{cases}2189.3=15a\\ 2024.3=280b\end{cases}$$

となり、これから

$$\begin{cases}a=145.75\\ b=\ \ 7.23\end{cases}$$

故に求める方程式は

$$y=145.75+7.23x\cdots\cdots\cdots\cdots(2)$$

となる。この(1)と(2)とは全く同一線である。蓋し(1)は一九一六年を零、(2)は一九二三年を零としたに過ぎず、從つて、例へば(1)から一九二三年のyを計算すれば、$y=95.14+7.23\times7=145.75$ となり、(2)に於てxを零とした一九二三年の $y=145.75+7.23\times0=145.75$ と一致する。他の年についても同様である。一般に $y=a+bx$ なる直線方程式のa（前例の 95.14 又は 145.75）は基準時（前例の一九一六年又は一九二三年）に於ける直線の高さであり、b（前例の 7.23）は一年毎に殖える高さである。（もし直線が $y=a-bx$ となれば、bは一年毎に減少する高さである。即ち斯かる傾向線は年と共に下降する線となる）。

グラフを一覽して發展傾向が直線的よりは寧ろ曲線的な場合には、それに應じて種々の形態の曲線を當嵌める必要があらう。併し曲線の種類は限りなく多く、こゝで一々説明する事は出來ぬ。以下では最も簡單な曲線、即ち拋物線 $y=a+bx+cx^2$ の當嵌めだけを問題としよう。この式を求める爲には a、b、c の三つを決定せねばならず、從つて

年次	X	ガソリン輸出額（單位百萬バレル） Y	XY	X^2	X^2Y	X^3	X^4
1915	0	2.7	0	0	0	0	0
1916	1	8.5	8.5	1	8.5	1	1
1917	2	9.9	19.8	4	39.6	8	16
1918	3	13.3	39.9	9	119.7	27	81
1919	4	8.9	35.6	16	142.4	94	256
1920	5	15.3	76.5	25	382.5	125	625
1921	6	12.7	76.2	36	457.2	216	1296
1922	7	13.8	46.6	49	676.2	343	2401
1923	8	20.1	160.8	64	1286.4	512	4096
1924	9	23.3	254.7	81	2292.3	729	6561
1925	10	30.6	306.0	100	3060.0	1000	10000
1926	11	42.5	467.5	121	5142.5	1331	14641
1927	12	44.3	531.6	144	6379.2	1728	20736
1928	13	52.9	687.7	169	8940.1	2197	28561
1929	14	62.1	869.4	196	12171.6	2744	38416
1930	15	65.6	984.0	225	14760.0	3375	50625
	120	431.5	4614.8	1240	55858.2	14400	178312

三箇の聯立方程式を必要とする。然るに直線と同様、最小自乗法の原理から次の公式が得られる。

$$\begin{cases}\Sigma(Y)=Na+b\Sigma(X)+c\Sigma(X^2)\\ \Sigma(XY)=a\Sigma(X)+b\Sigma(X^2)+c\Sigma(X^3)\\ \Sigma(X^2Y)=a\Sigma(X^2)+b\Sigma(X^3)+c\Sigma(X^4)\end{cases}$$

これを上の一九一五年乃至一九三〇年（十六年間）の米國に於けるガソリン輸出量（單位百萬バレル）に適用して見よう。直線の場合と同じく、年次は先づ一九一五年を零とする。右式に代入すべきものは $\Sigma(Y)$、$\Sigma(XY)$、$\Sigma(X^2)$、$\Sigma(X^2Y)$、$\Sigma(X^3)$、$\Sigma(X^4)$ である（Nは16である）。これらを表から計算し、右式に代入すれば

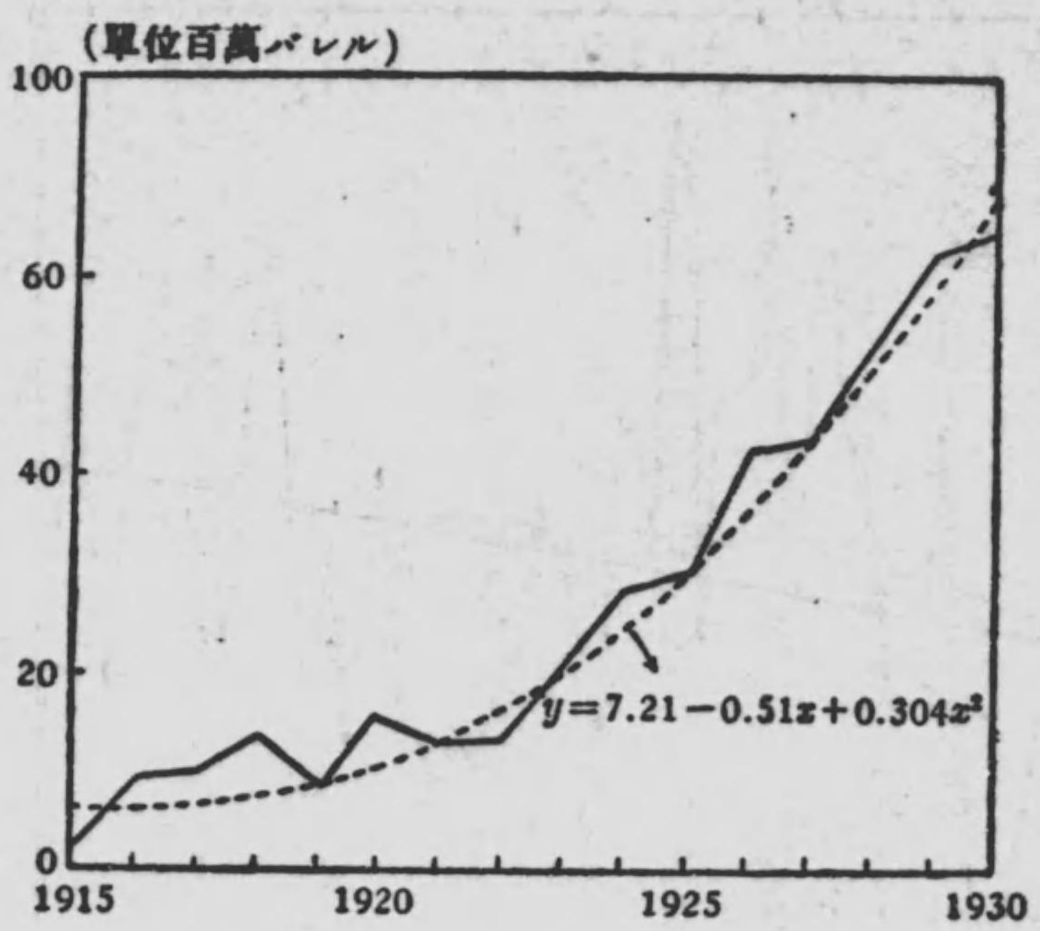
$$\begin{cases}431.5=16a+120b+1240c\\ 4614.8=120a+1240b+14400c\\ 55888.2=1240a+14400b+178312c\end{cases}$$

これを解いて

$$\begin{cases}a=7.21\\ b=-0.51\\ c=0.304\end{cases}$$

となるから、これを $y=a+bx+cx^2$ に代入すれば、求める拋物線の方程式は

$$y=7.21-0.51x+0.304x^2$$

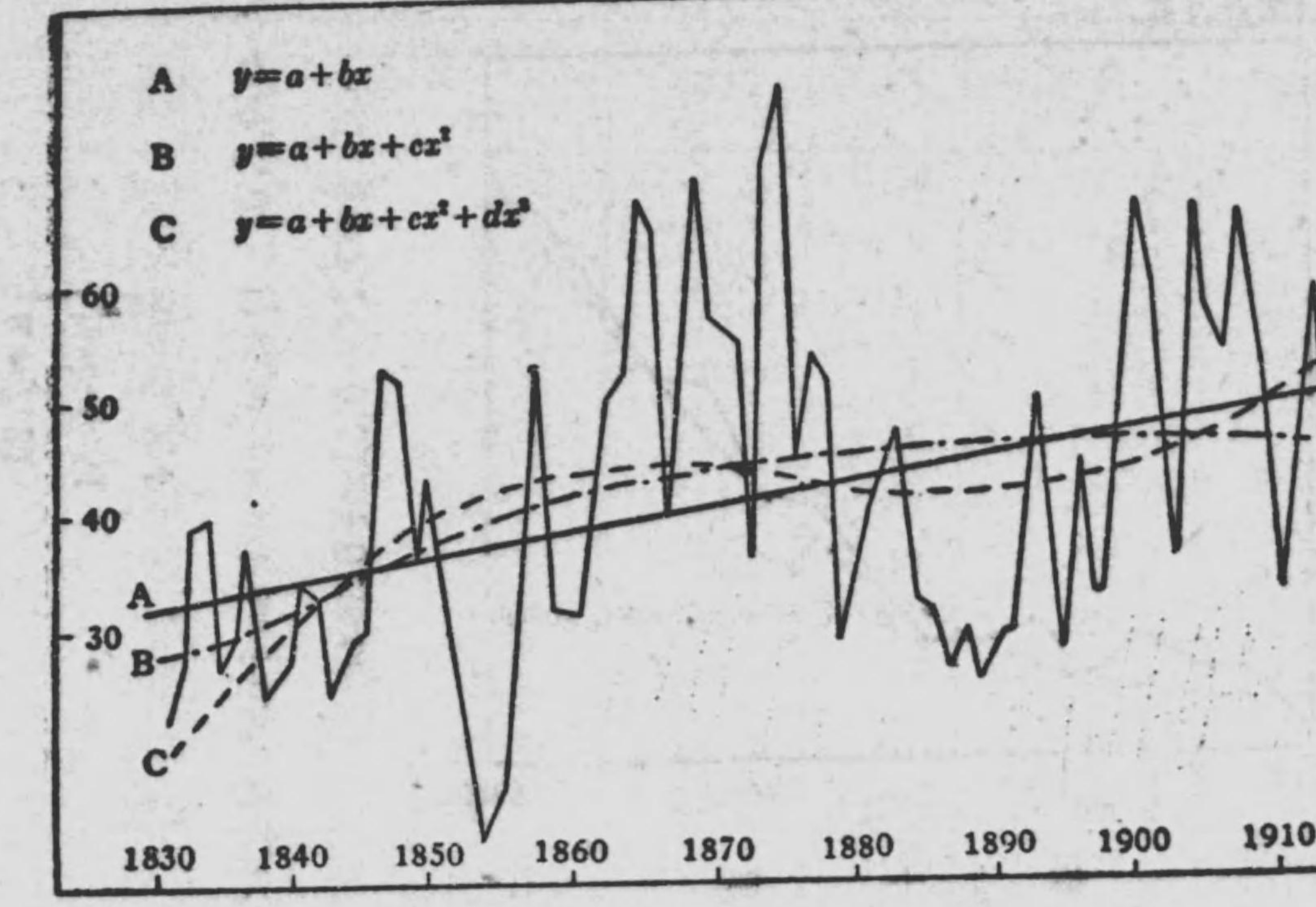


第十圖

となる。この x に順次に0、1、2、3………15を代入すれば、一九一五年から一九三〇年までの各年に於ける傾向線の高さが求められる。例へば一九一五年は $y=7.21-0.51\times0+0.304\times0^2=7.21$ であり、次年は $y=7.21-0.51\times1+0.304\times1^2=7.004$、次の一九一七年は $y=7.21-0.51\times2+0.304\times2^2=7.40$ である。直線の場合と異り、單に二時點の y の高さを決定した丈けでは曲線は描けない事に注意されたい。即ち各年につき y を計算せねばならないのである。斯くする事によつて、上の第十圖の拋物線（點線）が得られる。

註（い）より高次の曲線は一般には餘り用ひられない。併し例へば $y=ab^x$ の如き曲線は對數に書直せば比較的容易に描き得る。即ちその兩邊の對數をとれば $\log y=\log a+x\log b$ となり、直線方程式と同じ形になる。換言すれば、斯かる曲線は半對數圖表の上で直線となるのである。

人口統計に於て最も廣く利用されるゴムペルツ曲線（$y=ab^{c^x}$）又はパール・リード曲線$\left(y=d+\frac{b}{e^{-ax}+c}\right)$等については、此處では省略せざるを得ない。



第十一圖

第三節　長期傾向線の決定について

傾向値の決定に關する最も合理的な方法は上記の數學線の當嵌めであるが、實際にこれを適用する場合には屢〻大なる困難に遭遇する。それは與へられた系列に直線を當嵌むべきか、拋物線を當嵌むべきか、又はより高次の複雜な曲線を當嵌むべきかは、必ずしも客觀的に決定されるものではなく、主として個人的判斷によつて行はれるといふ事である。第十一圖は一八三〇年から一九一〇年までの佛國葡萄產額のグラフに直線、拋物線及び三次曲線を當嵌めたものであるが、その何れが最も理想的かは斷定が困難であらう。

一般に、高次の曲線を用ふるほど與へられた諸點に接近した傾向線を描き得るものであるが、併し傾向線とは定義上、與へられた時系列の一般的方向に外ならぬから、高次曲線による過度の近

似線は寧ろ不適當といはねばならぬ。斯くて高次曲線は除外するとしても、而も直線とすべきか又は拋物線とすべきかの如き問題は常に殘されてゐる。直線と拋物線とでは、その示す發展傾向は甚だ異るのであるから、計算に取りかかる前に、豫め充分の注意を拂はねばならぬ。

註　高次曲線は複雜な波狀を示すから、一般的方向が判別されぬのみか、「長期波動」なる特殊の景氣循環運動との差別が困難となる。長期波動說は一九一三年、丁抹のゲルデレンに創り、コンドラティフによつて一應の完成を見た新學說である。後者は一七八九年から一九二〇年に至る約百三十年間の英國卸賣物價指數につき、一七八九年から一八一四年までは上昇、一八四九年までは下降、一八七三年までは上昇、一八九六年までは下降、一九二〇年までは上昇の波動運動を發見した。氏はこれと類似の波動運動を幾多の經濟系列のうちに發見し、西半球に於ては十八世紀末以降二箇半の、卽ち平均五十年乃至六十年に及ぶ「長期波動」の存在を證明したのである。何故斯かる波動が生ずるかの理由は未だ充分に說明されてゐないが、大體に於て金の供給量の變動と一致する事は認められてゐるやうである。

斯かる波動の發見が、長期傾向と景氣循環との差別を困難ならしめたことは疑へない。本稿では、この長期波動の問題には觸れぬ。

第六章　時系列の解析　(二)

第一節　季節變動の意義

凡ゆる一切の運動のうち最も規則的に繰返へされるものは、恐らく天候なる文字に要約される自然現象であらう。溫度とか降雨量とかの所謂天候なるものは、大體に於てその土地の自然的狀態によつて決定されるもので、現在の人力を以てしては殆ど左右されざるものである。春夏秋冬の循環、雨期と乾燥期の循環なぞは毎年規則正しく繰返へされる。

そして是等こそ、農業漁業の如き自然的產業を根本的に決定する條件なのである。

毎年繰返へされる斯かる自然現象を本來の意味に於ける季節（Season）といふ。斯かる季節は二つの意味に於て吾人の生活に一定のリズムを發生せしめる。卽ち第一には、夏は氷を、冬は薪炭を、需要せしめるといふ風に、季節は直接に吾人の生活に規則的影響を與へる。第二には、季節は吾人生活の基本たる農業その他の原始產業を律するから、間接に吾人の生活に或る秩序を生ぜしめる。例へば收穫が秋に行はれる結果、毎年その頃には貨物輸送量が增大するとか、その取引の爲に特に多額の資金が流動するとかの如きこれである。

如上の自然的季節に對して、社會組織の上から發生する人爲的季節なるものがある。例へば一年を十二箇月に分つ

とか、七日を以て一週とするとかは、元來生活の便宜の爲に人が任意に定めたもので、格別自然的根據のあるものではないか、而もこれによつて吾人の日常生活が律せられてゐる事は疑へない。定期的の大祭や祝日も亦同樣である。是等は何れも規則正しく循環するものであるから、これ亦季節と呼んで差支へないのである。

蓋し斯かる人爲的季節は、例へば月末にその月の取引の決濟を行ふとか、日曜や祭日には仕事を休むとかの形に於て、吾人の生活に一定の習慣を來さしめるからであつて、その作用は上記の自然的季節と同樣、極めて規則正しいのである。

この自然的及び人爲的の二つの季節は相合して、社會生活に一定の季節的變動を發生せしめる。吾人が一般に季節變動といふ場合には、この綜合的結果を指すのであつて、從つて從來統計學の課題は、この綜合的季節變動を他の運動形態、即ち長期傾向、景氣變動（循環運動）及び不規則變動から如何にして分離せしめ得るかといふ事であつた。次に述べる各種の測定法は、何れも右の意味に季節變動を一體として測定せんとするものである。併し解析の最後的楷段は、一箇の綜合物をば其の不可分的分子にまで分析する事であるから、季節變動は最後には更に自然的及び人爲的の二つの變動に分析されねばならぬ。現在の統計的解析法を以て果して之を爲し得るや否やは、本章の終りに述べよう。

註 季節變動とは元來一年を以て一巡する循環運動を指すが、時にはより短期の——例へば一箇月、一週、又は一日を以て一巡する循環運動を指す事もある。こゝでは前者のみを取扱ふ。併し後者に於ても、その算出方法は全く同一である。

第二節 季節變動の測定

時系列に包含される既述の四運動形態は、系列によつて著しく組合せを異にする。自然現象に關する系列に於ては、吾人の言ふ意味の長期傾向は認められぬ。溫度や降雨量の如きこの例である。且つこれら系列に於ては、數年を週期とする循環運動が存在することの立證されてゐるものも少くないが、その程度は極めて精密な測定によつて判る位ひのもので、比較的短い期間については殆ど無視し得る程度のものである。(註) 故に斯かる自然的現象の變動は殆ど全く季節的及び不規則的の二變動だと言つてよからう。して見れば斯かる場合に季節變動を求めるには、單に不規則變動を除きさへすればよいのであつて、その操作は當然簡單なるべき筈である。次の月別平均法はこの目的に應ずるものである。

註 景氣理論の一つとして天體的景氣論なるものがある。或る天體現象に週期的運動があり、それが直接又は間接に地球に影響を與へる事によつて吾人の經濟生活に循環的起伏を發生せしめるといふのであつて、例へばジェヴォンス(Jevons)の太陽黑點說、ムーア(Moore)の金星相合說の如きその代表的なものである。詳しくは最近の拙稿「天體的景氣理論の二類型」(三田學會雜誌、昭和十二年四月號)を參照されたい。

反之、大部分の經濟系列に於ては、特殊の場合を除き、上記の四運動形態は必ず同時に存在してゐるのである。特殊の場合とは、例へば專賣法によつて規定された煙草、小賣價格の如きもので、これらは規定の改められぬ限り、常に同一を保ち、一切の變動を免れてゐるのである。併し固より斯かる經濟系列は異例に屬し、大部分の經濟系列に於

ては、四箇の變動が多かれ少かれ參加してゐるから、從つて季節變動を抽出する方法も當然複雜とならざるを得ない。十二箇月移動平均法や連環比率法はこの目的の爲に案出された方法である。

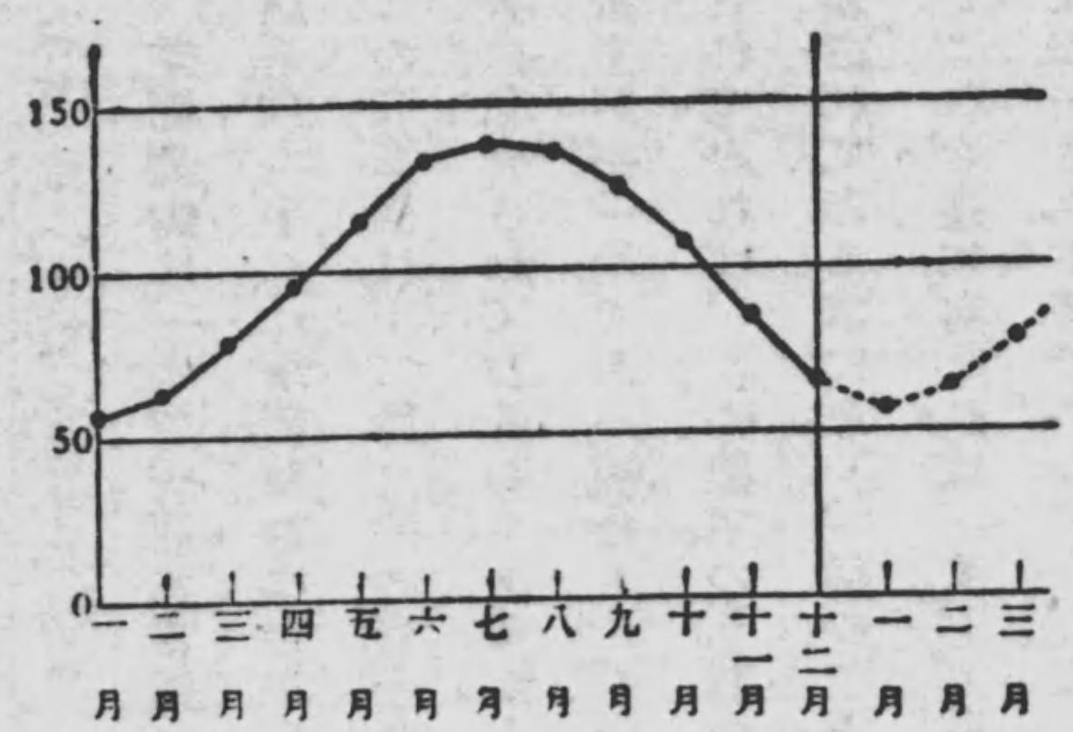
(A) 月別平均法

溫度とか降雨量とかの自然的現象に就ては、上述の如く長期傾向は認められず且つ循環運動も必ずしも明瞭ではないから、斯かる現象の統計系列に觀取される變動は、季節變動と不規則變動との綜合物と見て差支へない。故に若し不規則變動さへ除去し得るならば、季節變動は自ら現はれて來る筈である。いま紐育市の溫度表を一覽して見よう。同市では一月が最も寒い月とされてゐるが、併し毎年必ずそうだとは限らない。二月又は十二月の方が寒い年は屢、見受けられる。故に一月が最嚴寒の月だと言ふ爲には、相當長期間の平均を採る外はあるまい。平均とは、既に説明した通り、事物の正常狀態に外ならぬから、長期間内の一月の平均溫度が他の月のそれよりも低ければ、一月を以て最嚴寒の月と認めてよからう。月別平均法とは、月別統計を各月毎に平均する事によつて、毎月の正常値を決定する方法を言ひ、これを指數に換算すれば、月の經過に伴ふ季節變動の大さ即ち季節指數を求め得るのである。

紐育市溫度の季節指數

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月	平均
1917	32	28	39	47	53	68	74	75	63	52	41	25	50
1918	22	30	41	30	64	66	73	75	63	59	46	39	52
1919	35	35	42	49	61	70	74	70	66	58	44	30	52
1920	24	29	41	48	58	68	72	73	67	60	44	38	52
1921	34	35	48	55	60	70	76	71	71	56	45	34	55
1922	29	34	41	51	64	71	73	71	67	58	45	34	53
1923	31	27	37	49	59	72	72	71	67	56	44	42	52
平均	30	31	41	50	60	69	73	72	66	57	44	35	52
指數	58	62	78	95	114	132	139	137	126	109	84	66	100

右の一九一七年から一九二三年に至る七年間の紐育市華氏溫度表を見るに、一九一七年の如く十二月が最も低く、八月が最も高いやうな年もあるが、平均を求めれば一月が三十度で最も低く、七月が七十三度で最も高い。これは平均によつて極端な値が相殺され、最も自然的正常的數値が現はれるからで、從つて右表の各月の平均溫度をば典型的溫度と見て差支へなからう。この平均溫度は季節の影響を直接且つ具體的に示すもので、これによつて紐育市溫度の季節變動を知り得るわけである。併し季節變動の大さは成るべく指數に換算して置く方が比例に便利である。右表の最後の欄は各月の平均溫度を總平均(五十二度)に對する百分比に換算したもので、これを季節指數と言ふのである。これを求めるには、例へば一月については、$52:30=100:x$ 即ち $\frac{30}{52}\times100=58$ なる運算を施せばよい。他の月についても同樣である。斯くすれば指數の合計は一二〇〇、即ち月平均一〇〇となり、或る月の指數が一〇〇よりどれだけ大なるか小なるかを見れば、年平均溫度からの隔りの程度が容易に見分けられよう。更にこれを圖示すれば、一層明瞭に季節變動の樣態を知り得るであらう(第十二圖參照)。



第十二圖

（B）十二箇月移動平均法

季節變動は一年で一巡する運動であるから、與へられた材料が一年の合計又は平均によつて示されてゐる場合には、該系列の中には季節變動は存在しない筈である。例へば或る商店の一年間の賣上合計が五萬圓と發表されたとすれば、この五萬圓は十二箇月間の賣上合計であり、從つてそのうちには季節的に賣上の增加する時期も、その反對の時期も共に包含されてゐるのであるから、この單一の數字から季節變動を抽出する事は不可能である。この事は與へられた數字が年平均である場合も全く同樣である。例へば或る年の東京市場金利統計に於て商業手形金利が一錢三厘と發表されてゐれば、この數字のうちには金利の高い年末も低い夏期も平均されてゐるのであるから、季節の影響は全く相殺されてゐる事になり、從つて季節變動を抽出する事は不可能なのである。

これを逆に考へれば、もし月別統計が與へられてゐる場合でも十二箇月分を合計するか又は平均すれば、季節變動は消去される筈である。そして斯く季節變動が消去されたならば、これと原系列（即ち季節變動を含む月別統計）との差又は比を求めるならば、季節變動そのものが現はれて來ねばならぬ。對十二箇月移動平均比率法（Ratio to 12-Months Moving Average Method）とは、斯かる根據の上に作られた方法である。

その方法の順序は、〔第一〕連續する各月の十二箇月平均を順次に求める。即ち一月から始まる數字であれば、先づ一月から十二月までの平均を求め、次には一箇月をづらして二月から十二箇月即ち次年の一月までの平均を求め、以下同樣にして最後の月に至るまでこの計算を繰返へす。〔第二〕この一月から十二月までの平均を七月の値とし、次の二月から次年の一月までの平均を八月の値とし、以下同樣にする。（註）〔第三〕移動平均が求められたならば、順次に原數をその該當する移動平均値で割つて、原數の平均値に對する比率を求める。例へば七月の原數をその月の平均値で割り、八月の原數をその月の平均値で割り、以下同樣にして最後までこの計算を行ふのである。〔第四〕斯く求められた移動平均値は、その性質上、最初と最後の各六箇月づつの値は得られないから、例へば十年間に亙る材料が與へられてゐる場合にも、移動平均値は百二十箇ではなく百〇八箇、即ち各月何れも九箇づつの平均値を得ることになる。この九箇の一月の値、九箇の二月の値………九箇の十二月の値から、各〻その月の代表値と認められるものを求め

A　米鷄卵價格（單位、仙）

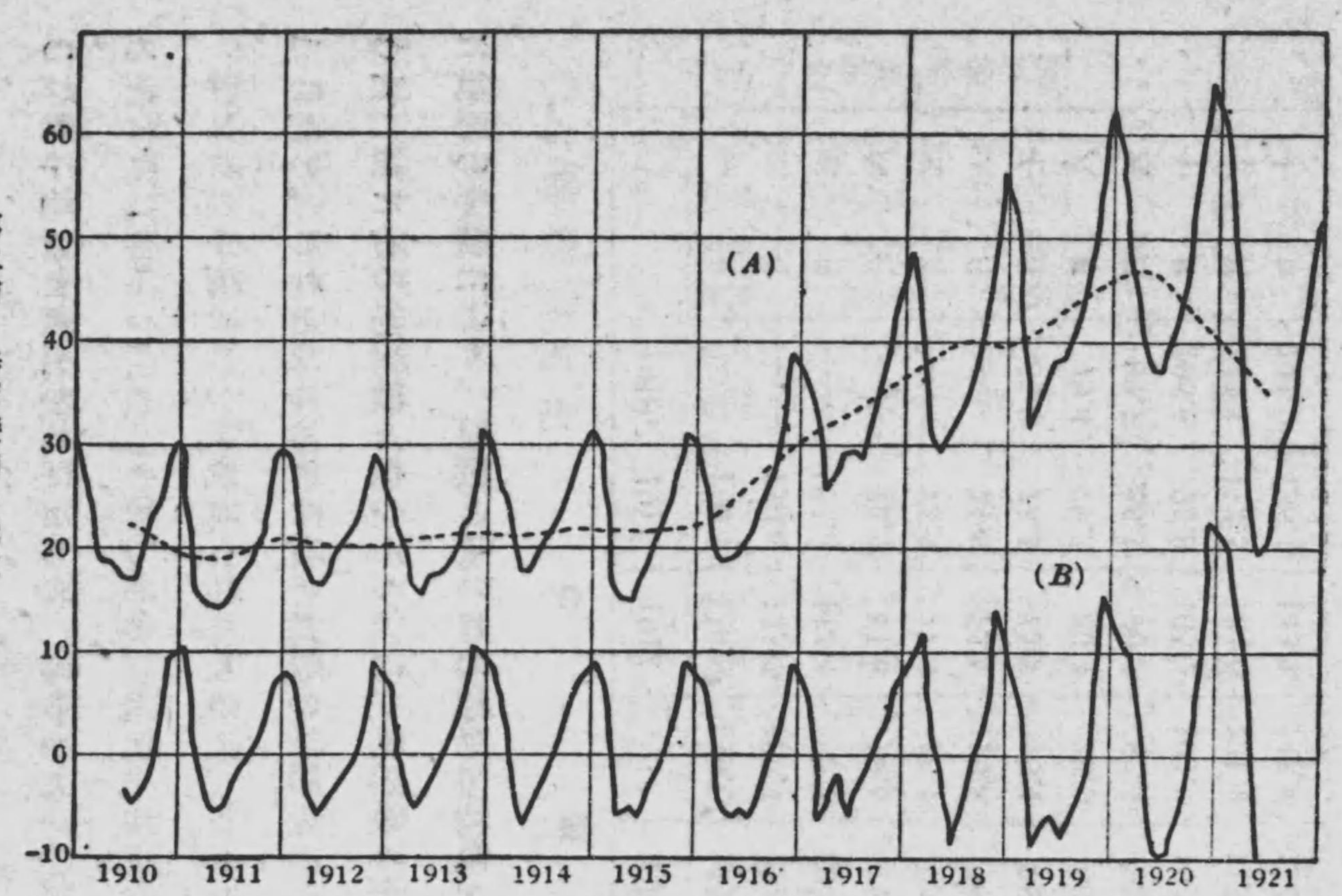
	1910	1911	1912	1913	1914	1915	1916	1917	1918	1919	1920	1921
一月	30.5	30.4	29.5	26.8	30.7	31.6	30.6	37.7	46.3	57.2	64.8	61.1
二月	23.9	22.1	29.1	22.8	28.4	29.2	26.8	35.8	49.4	48.3	56.9	49.6
三月	22.9	16.5	24.5	19.4	24.2	21.3	21.2	33.8	40.4	33.1	46.6	29.2
四月	18.6	14.9	17.8	16.4	17.6	16.6	17.9	25.9	31.2	34.3	38.8	20.4
五月	18.6	14.7	17.1	16.1	16.8	17.1	18.1	30.0	31.0	36.8	37.4	20.2
六月	18.3	14.5	16.7	16.9	17.3	16.6	19.0	31.1	29.8	38.6	37.0	19.4
七月	18.2	14.2	16.7	17.0	17.6	16.8	19.7	28.3	30.7	36.8	36.7	22.0
八月	17.6	15.5	17.4	17.2	18.2	17.0	20.7	29.8	34.4	39.3	40.0	26.6
九月	19.4	17.4	19.1	19.5	21.0	18.7	23.3	32.2	36.4	41.0	44.2	30.4
十月	22.4	20.0	22.0	23.4	23.5	22.3	28.1	37.4	41.6	44.7	50.1	34.2
十一月	25.3	23.5	25.9	27.4	25.3	26.3	32.2	39.4	47.2	54.0	56.9	44.2
十二月	29.0	28.7	29.7	23.0	26.7	30.6	38.1	43.3	55.0	61.9	65.0	51.1
平均	22.5	19.4	22.1	21.3	22.5	22.0	24.6	33.8	39.5	43.8	47.9	34.0

る。これには九箇の値の平均値を求めればよいのであつて、この場合算術平均でも、中位數でも、又は中央數箇の算術平均でも差支へがない。〔第五〕斯くて得られた一月乃至十二月の十二箇の代表値は各月の季節變動を示すものであるが、便宜上この十二箇の比率をその平均で割つて、指數の合計が一二〇〇となるやう換算する。この手續をミル氏の資料に據つて實際に説明しよう。この資料は一九一〇年から一九二一年までの十二年間の鷄卵一ダース卸賣價格(單位、仙)である(A表)。

註 一月から十二月までを平均した値は、實は六月と七月との中間に置かる可きものである。七月に置けば明かに半箇月の誤差が

B 表 (十二箇月移動平均)

	1910	1911	1912	1913	1914	1915	1916	1917	1918	1919	1920	1921
一月	——	20.25	21.26	20.78	22.74	22.29	22.24	30.06	36.72	41.39	46.60	40.47
二月	——	19.99	21.45	20.79	22.80	22.20	22.51	30.80	37.01	41.86	46.63	39.30
三月	——	19.82	21.60	20.79	22.91	22.06	22.86	31,59	37.33	42.25	46.80	38.16
四月	——	19.64	21.75	20.87	22.97	21.91	23.29	32.39	37.64	42 57	47.15	36.93
五月	——	19.46	21.94	20.99	22.89	21.90	23.77	33.08	38.14	42.98	47.50	35.74
六月	——	19.38	22.08	21.19	22.66	21.98	24.33	33.59	38.96	43.56	47.75	34.63
七月	22.47	19.33	22.00	21.49	22.57	21.98	24.94	34.17	39.90	44.17	47.72	——
八月	22.18	19.58	21.63	21.88	22.64	21.84	25.60	35.10	40.31	44.84	47.26	——
九月	21.63	20.20	21.16	22.32	22.55	21.74	26.51	35.93	39.96	45.76	46.24	——
十月	21.21	20.66	20.89	22.57	22.39	21.78	27.36	36.43	39.79	46.50	44.74	——
十一月	20.89	20.88	20.79	22.65	22.36	21.88	28.20	36.69	40.17	46.71	43.26	——
十二月	20.57	21.07	20.76	22.69	22.35	22.02	29.20	36.68	40.77	46.67	41.81	——



第 十 三 圖

生ずる。これを修正するには、斯く七月に置いた値と八月に置いた値とを平均してこれを七月の値とすればよい。他の月についても同様である。要するに一度十二箇月移動平均によつて求めた値を更に二箇月移動平均によつて訂正するのである。此處では半箇月の誤差は無視し得るものと假定して、敢へて斯かる嚴密な手續を取らなかつた。

A表を圖示すれば第十三圖(A)の太線となり、毎年冬は高く、夏は低く、顯著な季節變動を示してゐる。この材料につき上記の手續を施したものが、B表である。即ち、一九一〇年七月の値二二・四七は、一九一〇年一月から同年十二月までの平均値であり、八月の二二・一八は二月から一九一一年までの平均値であり、以下總て同様である。既述の如く、この移動平均は一九二一年六月で終つてゐる事に注意されたい。是等の移動平均値を相互に連結したものが、第十三圖(A)の點線である。

この移動平均値は季節變動の消去された値である事は既に説明した。故に季節變動を求めるには原數からそれぞれ

の移動平均値を差引かねばならぬ。即ち一九一〇年七月の季節變動の大さは $18.2-22.47=-4.3$ であり、同年八月のそれは $17.6-22.18=-4.6$ である。表示すれば次のC表の數字となり、グラフに描けば第十三圖(B)となる。

併し右の手續によつて求められたものは、各年の各月に於ける實際の季節變動であつて、從つて總括的な季節指數ではない。これを求めるには右の實際の季節變動を更に平均したものでなくてはならぬ。然るにこれを平均するに當つて、單に右の絶對數を採つたのでは、絶對數の大なる項の影響が過度に強く現はれて、適當な平均を求め難い。故に第三の手續により、原數とその月の移動平均値との比率を求める必要が起つて來るのである。次のC表は、これを

C　表

	1910	1911	1912	1913	1914	1915	1916	1917	1918	1919	1920	1921
一月	—	150.1	138.8	129.0	135.0	141.8	137.6	125.4	126.1	138.2	139.1	151.0
二月	—	110.6	135.7	109.7	124.6	131.5	119.1	116.2	133.5	115.4	122.0	126.2
三月	—	83.2	113.4	93.3	105.6	96.6	92.7	107.0	108.2	78.3	99.6	76.5
四月	—	75.9	81.8	78.6	76.6	75.8	76.9	80.0	82.9	80.6	82.3	55.2
五月	—	75.5	77.9	76.7	73.4	78.1	76.1	90.7	81.3	85.6	78.7	56.5
六月	—	74.8	75.6	79.8	76.3	75.5	78.1	92.6	76.5	88.6	77.5	56.0
七月	81.0	73.5	75.9	79.1	78.0	76.4	79.0	82.8	76.9	83.3	76.9	—
八月	79.4	79.2	80.4	78.6	80.4	77.8	80.9	84.9	85.3	87.6	84.6	—
九月	89.7	86.1	90.3	87.4	93.1	86.0	87.9	92.4	91.1	89.6	95.6	—
十月	105.6	96.8	105.3	103.7	105.0	102.4	102.7	102.7	104.5	96.1	112.0	—
十一月	121.1	112.5	124.6	121.1	113.1	120.2	114.2	107.4	117.5	115.6	131.5	—
十二月	141.0	136.2	143.1	145.4	132.9	139.0	130.5	118.1	134.9	132.6	155.5	—

一纏めにしたもので、例へば一九一〇年七月の八一・〇は、原數の一八・二をその月の移動平均値二二・四七で割り、それを百倍して百分比としたものである。

この百分比の表から第五の手續によつて各月の代表値を求めるのであるが、多くの場合に中位數を採る。即ち例へば一月の中位數は、一月の百分比十一箇を大さの順に配列し、中央に當る項の百分比即ち一三八・二である。最後に第六の手續により、この十二箇月の十二箇の中位數の平均（この例では九九・六となる）で各月の中位數を割つて百倍すれば、一箇月平均が一〇〇、即ち總和が一二〇〇となる。これが求める季節指數である。この第五及び六の結果を一表に纏めれば、上の通りである。

月次	中位數	季節指數
一月	138.2	138.7
二月	122.0	122.5
三月	96.6	97.0
四月	78.6	78.9
五月	77.9	78.2
六月	76.5	76.8
七月	78.0	78.3
八月	80.4	80.7
九月	89.7	90.1
十月	103.7	104.1
十一月	117.5	118.0
十二月	136.2	136.7
平均	99.6	100.0

(C) 連環比率法

季節變動は主として月から月への變動と見られるから、與へられた月別資料をそのまゝ使用する代りに、先づ全期間に亙つて對前月比率（直前の月に對する比率）に換算し、これに基いて季節指數を算出する方が正しいといふ說がある。連環比率法（Link-relative Method）とは、ハーヴァード大學教授パーソンズ氏（W. Persons）が右主旨に基いて考案したもので、精密な季節變動測定を必要とする場合にはこの方法に據る外はない。

月別資料が與へられたとき、(一)各月に連環比率を算出する。連環比率とは或月の値をその直前の月の値で割つたもの、即ち對前月比を連續的に配列したものをいふ。これによつて各月の値はその前月の値に對する百分比として示さ

れる。(二)是等の連環比率表から、各月毎に代表値（中位數）を求める。(三)是等中位數を連鎖比率（Chain-relative）に書改める　連鎖比率とは連環比率の總ての中位數をば、一月を基準とした百分比に換算する事で、この爲には一月の中位數を一〇〇と置き、この一〇〇に二月の中位數を乘じたものを二月の連鎖比率とし、この二月の連鎖比率に三月の中位數を乘じたものを三月の連鎖比率とし、以下同樣にして十二月までの連鎖比率を求め得るのである。併し十二月の連鎖比率を求めた後、更にこれに一月の中位數を乘じて、別に一月の連鎖比率を求めて置く。この最後の一月の連鎖比率は、もし與へられた系列に長期傾向が存在しない場合には、最初の一月の連鎖比率（卽ち一〇〇）と一致するが、經濟系列には殆ど必ず上昇又は下降の長期傾向が介在するから、この二箇の連鎖比率は一致しないのが普通で、最後の一月の連鎖比率は一〇〇よりも大きいか又は小さいかを常とする。故に(四)この差を修正して兩者が一致するやう或る手續を施す。(五)最後にこの修正された連鎖比率の合計が一二〇〇となるやう、各月の比率をその平均で割れば、求める季節指數が得られる。次にこれを實例について說明しよう。

　次のA表は一九一四年から一九二九年に至る米國に於ける無煙炭生産額月別統計から作製した連環比率の表である（この資料は Arkin & Colton: An Outline of Statistical Methods から採つた。尤も計算の最後の部分は私が自分で行つた）。猶ほこの表は各月の連環比率を示したもので、生産額そのものでない事に注意されたい。例へば、一九一四年二月の八八・二六は二月の生産額はその前月たる一月の八八・二六％だつたといふ意味である。同年の一月の分が缺けてゐるのは、その前月たる一九一三年十二月の生産額が與へられてゐないために、計算出來なかつたからである。

A表　連環比率（米國に於ける無煙炭生産額）

年次	一月	二月	三月	四月	五月	六月	七月	八月	九月	十月	十一月	十二月
1914		88.26	128.16	51.94	120.92	110.02	109.23	110.03	103.36	96.59	88.59	113.39
1915	99.29	78.00	108.46	94.25	103.24	109.76	104.74	107.28	107.34	107.91	101.22	102.39
1916	101.70	97.00	96.99	76.78	115.37	97.27	109.98	112.04	98.59	109.44	100.27	98.15
1917	108.78	86.20	115.77	87.42	112.52	99.43	98.87	192.33	95.23	107.16	98.66	92.35
1918	85.89	103.67	109.89	95.70	109.57	101.39	107.49	100.25	92.87	102.19	83.94	91.53
1919	105.00	76.08	106.82	95.39	116.75	98.69	115.23	100.41	110.35	118.65	33.23	195.90
1920	133.59	82.53	116.54	81.00	102.79	115.71	99.76	108.65	100.54	106.05	98.66	101.29
1921	77.27	76.60	98.51	90.66	120.99	101.70	89.64	113.66	101.64	124.59	82.37	85.98
1922	123.00	108.99	122.41	31.46	128.58	109.95	76.19	152.05	158.67	110.06	100.36	102.64
1923	108.01	84.04	110.00	91.94	107.26	98.72	99.21	108.29	94.58	106.42	87.27	92.82
1924	127.33	90.08	87.29	73.70	106.08	97.46	106.01	107.71	117.98	114.23	87.97	109.90
1925	111.61	75.08	96.52	89.55	105.28	104.76	106.59	113.39	104.32	113.64	95.45	104.00
1926	101.31	86.72	99.05	86.88	97.46	107.51	103.51	106.64	105.66	111.47	109.38	96.38
1927	99.09	93.01	113.68	57.65	102.37	102.88	92.11	123.96	100.53	104.96	92.33	101.66
1928	107.46	93.53	106.31	73.22	113.76	98.20	100.89	113.31	100.46	121.94	91.42	94.22
1929	120.19	91.87	83.24	93.75	108.91	94.77	106.74	108.01	101.42	115.10	89.16	101.12

　次に各月の代表値を決定する爲に、中位數を求める。一月に關する十五箇の連環比率を大さの順に配列して中央の項の値を見れば一〇七・五であるから、これを一月の代表値とする。同樣にして二月、三月……十二月の代表値を決定する（但しこの例では二月以降の各月の連環比率は何れも十六箇づつであるから、中位數は中央二項の算術平均で

一月	107.5
二月	87.5
三月	102.6
四月	87.1
五月	109.2
六月	101.6
七月	102.3
八月	108.5
九月	101.5
十月	109.0
十一月	91.9
十二月	101.2

ある。例へば二月の中位數は八八・二六と八六・七二の算術平均即ち八七・五である)。斯く求めた各月の代表値を表示すれば、上表の如くである。

次で、㈢の手續により、上表から連鎖比率を作製する。即ち、一月の値を一〇〇と置き、これに八七・五を乗じた八七・五（勿論一〇〇で割る）が二月の連鎖比率であり、この八七・五に一〇二・六を乗じた八九・八が三月の連鎖比率であり、以下同様である。即ち次表の如くである。

一月	100.0
二月	87.5
三月	89.8
四月	87.3
五月	95.3
六月	96.8
七月	99.1
八月	107.5
九月	109.1
十月	118.9
十一月	109.3
十二月	110.6
一月	118.9

この表に於て最後に再び一月の連鎖比率を算出したのは、既に説明した通り、これによつて長期傾向値の有無を見る爲である。その値が一〇〇となれば、長期傾向は存在しない事を意味する。表では一一八・九となり（この値は十二月の連鎖比率一一〇・六に一月の連環比率一〇七・五を乗じたものである）、一八・九だけの差を生じてゐる。この差を修正する爲には、二種の方法が行はれてゐる。一はこの差を十二等分し（この差は一月から一月まで、即ち十二箇月間に生じたものだからである）、各月の連鎖比率から順次に月の數に應じて差引くのである。（但し最後の一月の値が一〇〇以下のときは、差の十二分一を順次加へて行く。）即ち右例では十二分一は $\frac{18.9}{12}=1.57$ であるから、二月に就ては $87.5-1.57=85.9$、三月に就ては $89.8-2\times1.57=86.6$ ……十二月に就ては $110.6-11\times1.57=93.2$ となり、最後の一月に就ては $118.9-12\times1.57=100$ となつて、最初の一月の連鎖比率一〇〇と一致する。斯く修正された各月の連鎖比率を表示すれば次の如くである。

一月	100.0
二月	85.9
三月	86.6
四月	83.6
五月	89.0
六月	88.9
七月	89.6
八月	96.4
九月	96.5
十月	104.7
十一月	93.5
十二月	93.2
一月	100.0

併し最初と最後の二箇の一月の値の差は、嚴密に言へば、右の如く均分的に修正さる可きではなく、寧ろ複利算の原則によつて修正さるべきである。蓋し十二箇月間に生じた差は、恰も預金に於ける元利合計と同様にして生じたと見るを至當とするからである。さて複利算に從へば利率を d、期間を n とすれば、n 年後の元來合計は $(1+d)^n$ である。いま一箇月の誤差は d、十二箇月なる期間は n と考へられるから、上式を適用すれば $118.9=(1+d)^{12}$ となり、これを解いて $d=0.00626$ が得られる。故に $\frac{118.9}{(1+d)^{12}}$ とすれば一〇〇となつて、最初の一月の連鎖比率と一致する。この修正法は他の月に就ても同樣であつて、例へば二月に就ては $\frac{87.5}{(1+d)}$、三月に就ては $\frac{89.8}{(1+d)^2}$ とすればよい。一般に n ケ月目の連鎖比率 (C_n) の修正は $\frac{C_n}{(1+d)^{n-1}}$ である。これは對數に書改めれば $\log C_n-(n-1)\log(1+d)$ となるから、下の如き計算表（B表）を作れば最も便利である。最後の欄の「修正された連鎖比率」は季節變動の程度を示

B　表

	連環比數の中位數	連鎖比率 (C)	$\log C$	$(n-1)\times\log(1+d)$	$\log C-(n-1)\times\log(1+d)$	修正されたる連鎖比率
一月	107.5	100.0	2.00000	0.00000	2.00000	100.0
二月	87.5	87.5	1.94201	0.00626	1.93575	86.2
三月	102.6	89.8	1.95328	0.01253	1.94075	87.2
四月	87.2	78.3	1.89376	0.01880	1.87456	74.9
五月	109.2	85.5	1.93195	0.02506	1.90689	80.6
六月	101.6	86.9	1.93902	0.03133	1.90829	80.9
七月	102.3	88.9	1.94890	0.03760	1.91130	81.5
八月	108.5	106.4	2.02694	0.04387	1.98307	96.1
九月	101.5	108.0	2.03342	0.05013	1.98329	96.2
十月	109.0	117.7	2.07078	0.05639	2.01439	103.3
十一月	91.9	108.2	2.03423	0.06266	1.97157	93.6
十二月	101.2	109.5	2.03941	0.06892	1.97049	93.4
一月	107.5	118.9	2.07518	0.07518	2.00000	—
						$Ma=90$

すものであるが、月別平均法に於て說明した通り、これら各月の連鎖比率をその平均九〇で割れば、月平均一〇〇、即ち年合計一二〇〇となり、十二箇月間の變動狀態を見るのに甚だ便利となる。この計算を行へば（例へば一月は $\frac{100.0}{90.0}\times100=111$　二月は $\frac{86.2}{90.0}\times100=96$ の如くする）上の結果が得られる。これが求める季節指數である。

一月	111
二月	96
三月	97
四月	83
五月	90
六月	90
七月	91
八月	107
九月	107
十月	115
十一月	104
十二月	104

第三節　季節指數の妥當性

經濟系列の解析に於て季節變動を測定せんとする場合、最も廣く採用される方法は連環比率法である。殊に最も精微な解析を必要とする景氣豫測に於て殆ど普遍的にこの方法が用ひられてゐる。併し十二箇月移動平均法も亦、その合理的な事は決して連環比率法に劣るものではない。

その何れの方法に據るにせよ、季節指數の算出は、他の一切の統計的解析に於けると同じく、可成り長期に亙る材料から出發せねばならぬ。併し注意すべきは、材料が長期間に亙れば亙るほど、季節變動の振幅は小さくなる傾きある事である。これは月から月への變動は決して恆常的ではなく、年によつて相當の差あるを免れないから、多年に亙る材料に於ては各月の變動振幅が多少とも相殺される傾きがあるからである。故に同一現象に關する異る二時期の季節指數を比較する場合、例へばA商品の戰前及び戰後の二つの季節指數を比較する場合、もし材料の期間が兩者に於て大いに相違すれば、正確な比較は不可能となるであらう。

一度び算出された季節指數が幾何の妥當性を有するかは、對象が溫度や降雨量の如き所謂「季節」そのものであるか、又は氷や石炭の季節的價格の如き「季節の社會的影響」であるかによつて著しく異るであらう。前者は自然現象そのものであるから、時代が變化してもその指數は殆ど不變である。現在の東京市の溫度の季節指數は恐らく百年前にも、百年後にも妥當しよう。反之、季節による商品價格の如きは生產技術、配給組織又は貯藏法等が改良されゝば、次第にその季節的變動は少くなるであらう。固より經濟系列の季節變動は必ずしも時と共に減少するとは限らない。その反對のものが決して少くないのである。即ち斯かる系列について算出された季節指數の妥當性は比較的に少く、從つて絕えず新たな材料によつて新たな指數を求めて行かねばならない。

註　季節變動に於ける自然的要素と人爲的要素の分解。

季節變動の原因は既述の如く、時には氣候や天候の如き自然的現象であり、時には制度や習慣の如き人爲的現象である。經濟系列に於ける季節變動は多くの場合この二原因の合成的結果である。例へば二月の交通量の減少は、一部は氣候寒冷の、一部は曆日の少い結果であるが、氣候は自然的現象であり、曆日の多少は人爲的制度である。制度や習慣は時に可成り動搖的であるから、統計的に求めた季節指數を更に分解し、その幾何が自然的原因に基き、幾何が人爲的原因に基くかを知る事は、現象の眞の認識には必要な事である。（曆日の多少に基く相違は、各月の一日平均の値を利用して容易に修正し得る。）

大戰前十五年間のベルリン市場割引利率の季節指數は

一月	二月	三月	四月	五月	六月	七月	八月	九月	十月	十一月	十二月
93	86	103	88	94	99	86	90	105	110	119	127

である。秋季の指數の外に、各四半期末、即ち三月、六月、九月及び十二月の指數の概して大なるが認められる。これは同市に於ける決濟上の習慣に基くもので、從つて人爲的原因に基く季節變動である。反之、秋季に特に高まるのは、農產物出廻による結果で、從つて自然的原因に基く季節變動である。この兩者は如何にして分離せしめられるか。

(1)最も單純な方法としては、季節指數の高まる各四半期末の各指數と、その隣接せる月の指數の差の平均を以て人爲的季節の大さと見るのである。右例に於ては約十三となるから、これが三月、六月、九月、十二月の人爲的季節指數となり、從つてこれらの月の原指數から各十三を差引けば自然的季節指數が得られる筈である。卽ち表示すれば次の如くである。

	一月	二月	三月	四月	五月	六月	七月	八月	九月	十月	十一月	十二月
人爲的季節	0	0	13	0	0	13	0	0	13	0	0	13
自然的季節	93	86	90	88	94	86	86	90	92	110	119	114

(2)併し右の方法では單に各四半期末の指數とその隣接せる月の指數を平均したのみであるから、必ずしも實際と合致しない。(例へば六月と五月の差は5に過ぎないのに、十二月と一月の差は30に達する)。これをより合理的に解決する爲に、著名な獨逸の統計學者パゥル・ロレンツ氏は次の方法を考案した。元來季節變動は一年内の連續的循環運動であるから、連續の概念に從へば、一一七頁に示したやうな月から月を直線で結んだ圖表は誤りである。換言すれば季節變動は各月の季節指數を縫つて走る近似的曲線によつて示されねばならぬ。斯くて前章の長期傾向線の當嵌めの手續が再び必要となるのである。唯だ異る點は、第一に長期傾向線は大體の方向を示す線であるから、直線又は拋物線の如き單純な線で足りるに反し、この場合には各月の指數に最も接近せる曲線が必要なのであるから、その方程式は必然複雜となり、第二に季節變動は每年繰返へされる規則的循環運動であるから、圖示法に於て說明した三角函數の方程式に據らねばならぬ事である。次の第十四圖（A）に於て、點線は單に各月の指數を直線で連結したものであり、曲線は右の原理によつて當嵌められた曲線であつて、その方程式は

$$y=100+12.1\sin(117.5^\circ+x)+9.27\sin(151^\circ+2x)+8.68\sin(101.5^\circ+4x)$$

となる。然らば上に問題とした季節變動の分析はこの曲線の分析に外ならない。

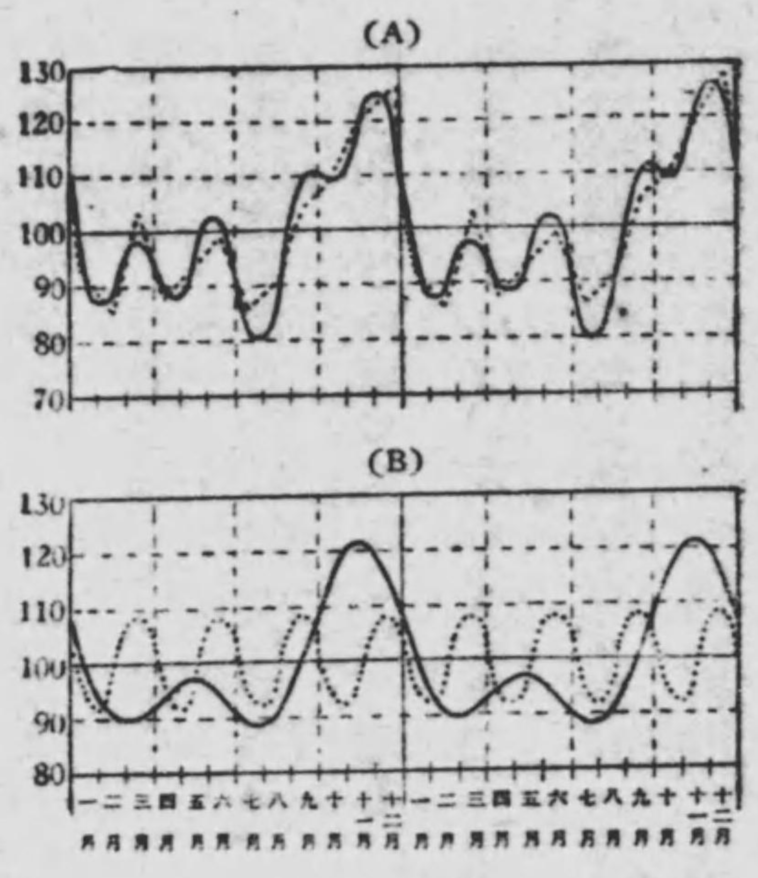
本例に於ける季節指數は各四半期末に於て人爲的原因によつて高まるが、いまこれら四時期に於ける決濟の習慣に相互に差なきものとすれば、この人爲的原因による運動は規則正しい四箇の波狀運動によつて示されねばならぬ。右の方程式に於てはこの波狀運動は

(1)　$y=100+8.68\sin(101.5^\circ+4x)$

によつて示される。然らばこれを除去した

(2)　$y=100+12.14\sin(117.5^\circ+x)+9.27\sin(151^\circ+2x)$

は當然自然的原因に基く季節變動を示す筈である。第十四圖(B)はこの二曲線を描いたもので、點線は(1)の方程式、卽ち人爲的原因による運動を、實線は(2)の方程式、卽ち自然的原因による運動を示すのである。



第十四圖

第四節　循環運動及び不規則變動の測定

始めに說明した通り、時系列に內在する變動形態は長期傾向、季節變動、循環運動及び不規則變動の四箇に分類される。上記の方法によつて最初の二形態、卽ち長期傾向と季節變動とが測定されたならば、この兩者をば原系列から除去する事によつて、殘餘の二形態、卽ち循環運動と不規則變動とを求め得る筈である。勿論この場合には、この二つの運動は未だ分離せず、相合して現はれて來るから、何等かの手段によつてこの兩者を分離する必要がある。この手續は後に述べる事として、先づ原系列から長期傾向と季節變動とを除去する方法を考へて見よう。いま或る系列の某年某月の數値は九六萬圓、長期傾向値は八五萬圓、季節指數は九三(%)とすれば、$96\div(85\times93)=121.5$(%)の算式

によつて、循環運動と不規則變動との大さを示す事が出來る。何となれば右式に於ける 85×93 は該月の正常値と認められるから（長期傾向値八五萬圓は長期間について見た該月の正常値であり、季節指數九五％は一年內に於ける該月の正常値であるから、從つてこの兩者を總括した 85×93 は該月の眞の正常値と認めざるを得ない）、この正常値と原系列との隔りは、當然殘餘の二運動形態の所產に外ならぬからである。即ち一般に或月の與へられた數字（原系列）を y、その月の長期傾向値と季節指數をそれぞれT及びSとすれば、その月の循環運動及び不規則變動の大さは $\frac{y}{T \times S}$ である。この計算を全期間の各月について行へばよい。次の假設例について見られたい。これは一九二一年一月から一九三一年十二月に至る十一年間の某商品生產額に關するものであるが、こゝでは單に最初と最後の二年分だけを擧げて置く。その傾向値は最小自乘法によつて當嵌められた直線の値であり、季節指數は連環比率法によつて測定されたものである（次表には單にその結果のみを記した）。

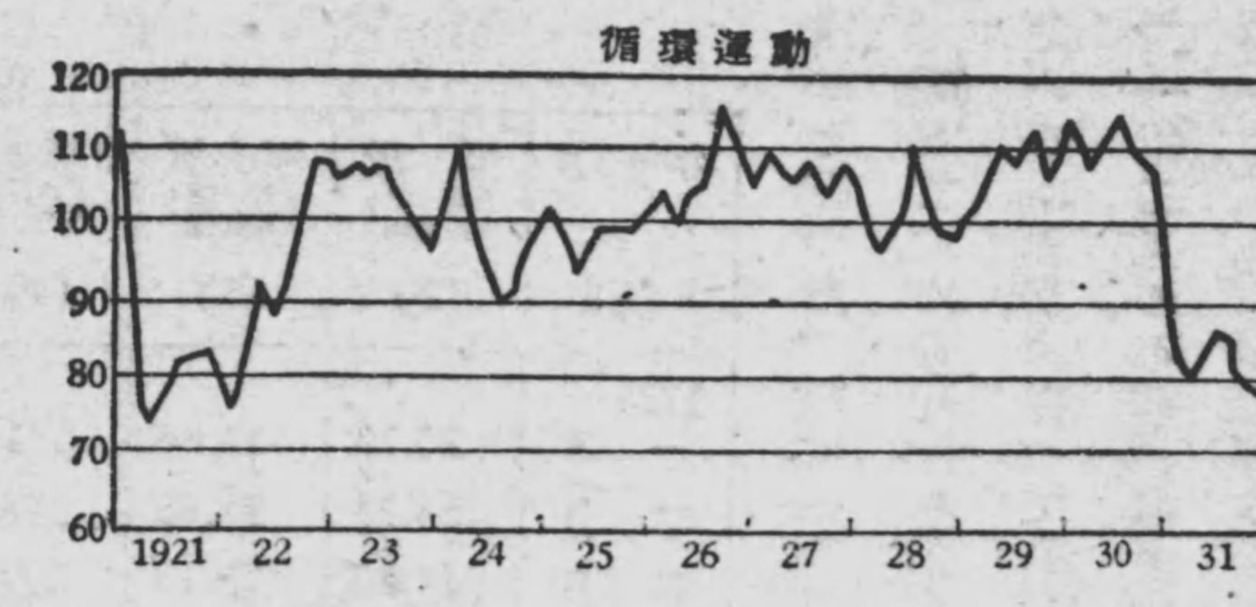
さて斯かる方法によつて原系列から長期傾向及び季節變動を除去し得るが、併し斯くして得た結果は循環運動と不規則變動との混合物であるから、次にこの兩者を分離せねばならぬ。然るに現在のところ、不規則變動を分離せしめる眞に適當な方法は無いのである。これは不規則變動なるものは全く偶然的理由によつて發生するもので、その生起の時期及び程度は時によつて著しく相違し、決して一率の取扱ひを許さないからである。併し既に說明した移動平均法を適用すれば、或る程度まで不規則變動を相互に相殺し、從つて眞の循環運動を抽出する事が出來る筈である。幾箇月の移動平均が適當なるかは一概に言へないが、一般に三箇月又は五箇月移動平均が用ひられるやうである。前例の結果に三箇月移動平均を適用すれば次の如くなる。但しこの場合、中心となる月の値に重きを置くため、二項式加

年　月 (1)	原數（單位百萬圓） (2)	傾向値（單位百萬圓） (3)	季節指數（％） (4)	正常値 (3)×(4) (5)	循環＋不規則（％） (2)÷(5) (6)
1921					
一　月	25.76	17.60	111.8	19.68	130.9
二　月	24.73	17.88	117.4	20.99	117.8
三　月	17.20	18.15	109.7	19.91	86.4
四　月	14.17	18.42	104.7	19.29	73.4
五　月	11.70	18.70	91.5	17.11	68.4
六　月	13.00	18.97	83.6	15.86	82.0
七　月	10.88	19.24	76.3	14.68	74.1
八　月	12.95	19.51	78.2	15.26	84.8
九　月	15.28	19.79	92.8	18.37	83.2
十　月	16.83	20.06	105.0	21.06	80.0
十一月	19.36	20.33	110.5	22.46	86.2
十二月	20.00	20.61	118.5	24.42	81.9
1931					
一　月	48.78	50.39	111.8	56.34	76.6
二　月	47.50	50.67	117.4	59.49	79.8
三　月	43.46	50.94	109.7	55.88	77.8
四　月	42.84	51.21	104.7	53.62	79.9
五　月	39.70	51.49	91.5	47.11	84.2
六　月	37.22	51.76	83.6	43.27	86.0
七　月	85.08	52.03	76.3	39.70	88.3
八　月	33.10	52.30	78.2	40.90	80.9
九　月	38.12	52.58	79.8	48.79	78.2
十　月	42.96	52.85	105.0	55.49	77.5
十一月	46.46	53.12	110.5	58.70	79.2
十二月	44.85	53.40	118.5	63.28	70.9

重法を採つた。故に一九二一年二月の平均値一一三・二は $\frac{130.9+2\times117.8+86.4}{4}$ として求められたものである（註）。

斯くて求められた移動平均値は、不規則變動の略、除去された値であるから、これを以て循環運動そのものと見てよいであらう。從つて若し不規則變動のみを知らんとするならば、右の移動平均値との差を採ればよい。例へば一九二一年二月の不規則變動は　$117.8-113.2=4.6$（％）である。

上述の如くして求められた十一年間の循環運動は次の如き圖表によつて示される。Y軸の一〇〇なる目盛は正常の



第十五圖

	(1)	(2)	三ケ月移動平均
1921			
一　月		130.9	―
二　月		117.8	113.2
三　月		86.4	91.0
四　月		73.5	75.5
五　月		86.4	73.1
六　月		82.0	76.6
七　月		74.1	78.8
八　月		84.9	81.8
九　月		83.2	82.8
十　月		79.9	82.8
十一月		86.2	83.6
十二月		81.9	83.5
1931			
一　月		86.6	89.7
二　月		79.8	81.0
三　月		77.8	78.8
四　月		79.9	80.5
五　月		84.3	83.6
六　月		86.0	86.2
七　月		88.4	85.9
八　月		80.9	82.1
九　月		78.1	78.6
十　月		77.4	78.0
十一月		79.1	76.4
十二月		70.9	―

狀態を示すものであるから、曲線が該線の上にあるか、下にあるかによつて、その時の狀態が正常狀態以上なるか以下なるかを判斷する事が出來る。循環運動とは所謂景氣變動に外ならぬから、景氣推移の樣態は物價・生産額・就業率等の主要系列について求められた斯かる圖表から知り得るのである。その詳細については別の「經濟統計」に於て説明するであらう。

註　二項式加重とは二項式展開に於ける各項の係數を以て、その項の重要さの程度と認める事である。即ち甲、乙、丙三項を平均する場合には、$(a+b)^2=a^2+2ab+b^2$ の係數を採ればよい。$a^2=1\times a^2$, $2ab=2\times ab$, $b^2=1\times b^2$ であるから、係數は 1, 2, 1 となり、從つて $\dfrac{甲\times1+乙\times2+丙\times1}{4}$ として平均値を求める。これを移動平均に適用すれば、右の平均値は乙の値となる。即ちこの方法によれば、乙の値を定める際に、乙に對して他の諸項の二倍の重要さを附與した事になる。加重法については、物價指數の章で詳しく説明する。

第七章　指　數

第一節　指數の意義と種類

物價指數、賃銀指數、生計費指數等の文字は、日常の新聞や雜誌に屢々散見するところであるが、これを正確に判讀する爲には、その構成に關する諸問題を一應心得て居らねばならぬ。蓋し是等は元來學問上の術語であつて、單に常識のみでは到底理解し難いからである。

指數とは既に「比率」を論じた際に説明した通り、比率の一種である。吾々が現象を數字的に把捉する目的は、實に比較に在る。然るに比較には差の比較と割合の比較とがある。甲は一尺、乙は二尺とすれば、「乙は甲よりも一尺長い」といふのは差の比較であり、「乙は甲の二倍の長さを持つ」といふのは割合の比較即ち比率である。然るに差を求める事は甚だ容易なのに對し、比率を求める事は時に困難を伴ひ、就中その一種たる指數を求める場合には特にその感が深いのである。殊にそれが物價指數とか生計費指數とかの特殊的內容を持つ場合には、單に算出上の數學的問題の外に、幾多の理論的及び實際的問題が起つて來る。

指數とは與へられた一統計系列の各項の大さを、基準項の大さに對する比率（普通%の形を採る）として示したものであるから、場所的系列、屬性的系列又は時系列の何れにも適用が出來る。即ち甲乙丙の三市の人口が夫々八萬、

六萬、五萬なりとすれば、乙市人口は甲市人口の $\frac{6}{8}\times100=75\%$、丙市人口は甲市人口の $\frac{5}{8}\times100=62.5\%$ に當るから、三市の人口比は 100, 75, 62.5 となる。これは場所的系列に於て甲市人口を基準とした三市の人口指數である。同樣に屬性的系列、例へば某工場の男女職工統計に於て、男工五百人、女工三百人なりとすれば、男工を基準とした兩者の指數は 100, 60 となる。併し指數が最も廣く利用されるのは時系列に於てゞある。統計的各種重要指數は殆ど常に、時の經過に伴ふ變化をば、或る時點（基準時點）の大さに對する%にて示したものと言つてよいのである。物價指數、生計費指數、生産量指數、賃銀指數等は總て時系列に於ける指數に外ならない。故に本章に於てはこの意味の指數のみを取扱ふ事にする。且つ右に掲げた各種の指數中、最も代表的な物價指數と生計費指數とは所謂「綜合指數」である。一系列の或る項を基準とし、他項をそれに對する%としたものは「單純指數」であつて、その作製は極めて容易である。然るに綜合指數とは幾多の系列を綜合し、これを一箇の指數に纏め上げたもので、その場合に理論的及び技術的に多くの問題が起るのである。この意味から以下の記述は特にこの二つ、就中物價指數に限定する。先づ物價指數から說かう。

第二節　物價指數の意義

吾々の扱ふ經濟の世界は、總て價格の世界である。或る現象が經濟學の對象となるのは吾人が該現象を價格と關聯せしめて考察するからであり、從つて價格と全く無關係な經濟現象なるものは、學問と無關係な學生といふが如き、考へ得られぬ概念である。然るに價格とは一定量の財貨と交換される貨幣量に外ならぬから、貨幣こそ經濟の世界の中心である。吾人の一切の經濟生活は、常に貨幣によつて律せられ、貨幣を媒介として行はれる。換言すれば貨幣こそ一切の經濟事象、一切の經濟價値の普遍的具體的尺度である。――この事は既に第一章に於て說明したつもりである。斯く貨幣が價格の基準とされ一般商品と交換される以上、貨幣自身が或る價値を持たねばならぬ事は明かである。長さを計る物指はそれ自身が長さを持たねばならぬし、容量を計る枡はそれ自身が容積を持たねばならぬと毫も異るものではない。

然るに貨幣の價値は、物指や枡と違つて、決して確定不變のものではない。同一種類の米一石が二十五圓のこともあるし、四十圓のこともあらう。同一量商品を購ふに少量の貨幣で足りるならば貨幣價値は大であり、反對に多量の貨幣を支拂はねばならぬならば貨幣價値は小さいわけであるから、同一商品の價格が絕えず變化する事實は、畢竟貨幣價値が絕えず變化する證據である。凡ゆる經濟事象を律する貨幣の價値が斯く變動するとすれば、この變動を正確に測定する事は、經濟事象の研究に於て不可缺の要件でなければならぬ。以下に說く物價指數は實にこの必要に應ぜんとするものである。

貨幣價値はそれと交換される商品の量によつて測られる。これは上に述べた所から明かであらう。こゝに言ふ商品とは一般に價格を有する一切の有形無形の事物又は行爲の總稱であつて、その範圍は極めて廣汎なものである。普通に言ふ商品は勿論、勤勞や權利の如きものも、それが貨幣と交換される以上、總て商品である。貞操の如き元來非經濟なものも、それが金錢によつて購はるゝ限りは、商品以外のものではあり得ない。貨幣はこれら無數の商品の何れとも交換される以上、單に一二の商品をとつてその價格から貨幣價値を判斷する事は許されない。十圓なる貨幣の價

値は、十圓で購ひうる一切の商品の量から判斷される。同時にして貨幣價値一般は、貨幣で購ひうる一切の商品の量から判斷される。

理論的には貨幣價値は斯くして決定されるが、併し實際にかゝる方法は行へるものではない。總ての商品を洩れなく枚擧する事は、星の數を數へるよりも困難であらう。そこで實際問題としては、事實存在する無數の商品につき嚴選を行つて、代表的な一部分について調査する外はない。事實無數の商品の中には、日常生活に關係の少いものや、價格決定の要因が充分合理的でないものが少くない。吾々の知らんとするのは一般の貨幣價値であるから、極めて特殊の用途に使用された場合の價値は全く無視しても大なる誤りは無からう。奢侈品、骨董品、權利等は勿論、一般に取引量の少いものや、勤勞等に支拂はれる場合は寧ろ除外するのが便利であらう。更に押進めれば、所謂完全な市場に於て合理的に決定される重要商品價格のみに限定するのが、最もよく客觀的に貨幣價値を知る所以となろう。

卽ち實際に貨幣價値を測定するには、二つの條件を必要とする。第一には經濟生活に關係の深い重要商品を選ぶこと。第二にはその價格が常に合理的且つ一般的に決定さるゝこと之である。この條件に合致するものは、調査さるべき商品としては結局米、麥、鐵、石炭等の主要商品であり、調査さるべき價格としては右商品の卸價格である。蓋し卸價格は客觀的一般的に決定される價格であつて、この點小賣價格とは趣きを異にするからである。

さて或る一定額の貨幣がこれら重要商品の何れをも購入する事が出來るならば、貨幣價値はこれら商品の價格の平均によつて判斷される筈である。そして商品價格は不斷に變化しつゝあるから、定時的にこれら平均を求めてゆけば、時の經過に伴ふ貨幣價値の變動を知る事が出來るであらう。價格の平均が高まれば、貨幣價値は下落した事になり、反對に價格の平均が下れば、貨幣價値は騰貴した事になる。この各時點に於ける價格平均を、一定時點に對する百分比として示したものが物價指數(Index Number of Prices)である。

物價指數の示すものは、右の如く一般貨幣價値であるが、併し吾人はそれ以外に特殊の貨幣價値を問題とする場合が少くない。吾々が日常使用する貨幣は、一般に小賣商品の購入に向けられるから、卸賣價格から作製された物價指數では、その價値を正確に測定する事は出來ぬ。何となれば小賣價格は、勿論卸價格に左右されるけれども、變動の樣態は著しく異るのである。故に個人的見地から見た貨幣價値は一般的貨幣價値とは別種のものであり、遙かに具體的現實的なものである。斯くて卸賣物價指數の外に、更に小賣物價指數を作製する必要が起つて來る。そして更にこの考へを徹底させれば、小賣商品中、特に生活に關係の深いものだけを選擇すれば、最も現實的な貨幣價値が測定される事が判らう。これを生計費指數といつて、收入乏しく、貨幣價値の變動によつて最も直接且つ深刻に生活の脅威を感ずる下層階級に關する理論的考察又は政策的施設は畢竟斯かる階級の生計費指數に立脚せざるを得ないのである。

物價指數作製に於ては理論的には算式の問題が中心であるが、實際的には如何なる商品を如何なる價格に於て調査すべきかの問題が中心となる。先づ後者から考察して見よう。

A、如何なる商品を調査すべきか

物價指數作製に使用すべき商品が所謂重要商品なるべき事は前述した。素より或る商品が重要か否かは決定困難であるが、取引額の大小に基準を求めれば略ゝ誤りなきを得るであらう。蓋し重要か否かは個人的見地からではなく、國民經濟の見地から決定さる可きものだからである。我國の代表的物價指數たる日銀東京卸賣指數が、取引年額約一

千萬圓以上の商品のみを採つてゐるのはこの適例である。勿論同一商品と雖も、その取引額は年々變動するから、若し例へば、取引年額一千萬圓を以て重要商品か否かを判別する嚴格な境界線とすれば、或る商品は年によつて重要商品となつたりならなかつたりする不便を生ずる。故に一千萬圓といふ樣な限界は決して嚴格なものではあり得ない。故に物價指數作製の計畫を樹てる際には、現在取引額の比較的尠いものでも、將來增進の見込みのあるものは採り、反之、現在取引額は多くとも將來の悲觀されるが如きものは寧ろ除外するのが普通である。

物價指數は貨幣價値の時の經過に伴ふ變化を知るを目的とするものであるから、一度決定した商品は中途に於て任意に變更さるべきではない。時に從つて商品種目が變れば、前後の比較が不可能となる事は言ふ迄もなからう。併し實際の問題としては、或る商品の持つ重要性は、人間の嗜好や生活の樣式の變化、又は新種目の出現等によつて著しく變つて行くものであるから、一度決定した商品種目の妥當性も素より永久的なものではあり得ない。日銀東京卸賣指數も昭和十二年末に種目其他の大改正を施し、一部を控除すると共に幾多の新種目を加へるに至つた。種目の實際に就ては後に述べよう。

商品種類の外に更にその數も重要な問題である。理論的には數は多いほど良いのが當然であるが、實際問題としては或る程度以上の數は單に調査及び計算を困難ならしめるのみで、結果は決してそれに比例するものではない。この問題の世界的權威たるフッシャー氏は、二十種以下では不正確であり、五十種前後ならば略〻充分であり、二百種以上は果して有利なりや否や疑問だと言つてゐる。米國の勞働省の物價指數の如く、三百種以上をとる例もあるが、多くは五六十乃至百二三十種である。要は比較的少數の最も代表的且つ一般的な商品を選擇すればよいのであつて、たとへ數は多くとも代表的ならざる或ひは偏つた商品をとつたのでは理想的指數は得られるものではない。

B、如何なる價格を調査すべきか

一般貨幣價値の測定に卸賣價格の適當な所以は、既に述べた通り、卸賣價格は專門業者間に決定さるゝ價格であり、最も合理的且つ一般的に商品の價値を、延いてそれと交換される貨幣の價値を示すからである。然るに卸賣價格にも種々の階段があり、製造業者、取引所、問屋、卸賣商、中央卸賣市場及び同業組合等の所謂「卸賣相場」があるから、その何れをとる可きかは一應問題とされねばならぬ。右のうち製造業者や問屋及び卸賣商の相場は對個人的性質が少からず含まれるし、又同業組合のそれは著しく統制された（卽ち自由ならざる）價格であるから、結局取引所と中央卸賣市場の相場がよい事にならう。蓋しこれら機關は所謂自由競爭の原則が最も露骨に作用し、組織は統一的、取引量は多量であるから、これらの相場を最も基準的な價格と見てよいのである。但し商品によつてはこれら機關では取引されぬものがあるから、それらに就ては他の機關の卸賣相場をとる必要があらう。

次に調査の時期であるが、卸賣相場は變動甚だしく、商品によつては一日に數回變化する事があるから、適當な方法を講ぜぬ限り、容易に決定されるものではない。普通に物價指數は毎月一回——時には二回又は毎週或ひは毎日の場合もあるが——發表されるから、一商品の價格は一箇月平均の價格たるを順當とする。これが爲には毎日の平均價格を求めてこれから月平均の價格を算出するのが最も合理的であるが、併し一日の變動の少いものについては、一日の或る時期の價格（多くは引値）を以て一日の平均價格に代用する場合が少くない。更に簡單な方法としては毎月の特定日、例へば中央日又は末日の相場を以て月平均に代へる場合もあり、事實廣く行はれてゐるのである。要するに

當該商品の一箇月間の價格を最も良く代表する價格を求めればよいのであるから、變動の少い價格については簡便法で充分な筈である。

物價指數の作製は相當の規模を必要とし、且つ永續的でなければならぬ上に、それ自身特定人の利益を目的とするものではないから、これが作製は官廳又は有力な研究所に俟つ外はない。我國でも日本銀行、商工省の如き官廳的機關、又は有力な經濟雜誌社（東洋經濟新報社、ダイヤモンド社）の事業となつてゐる。

第三節　綜合指數の算式

（一）基準　商品の種類が決定され、その價格が調査されたとして、次の問題はこれを如何にして物價指數の形に纏め上げるかといふ事である。物價指數は時の經過に伴ふ物價の變化（卽ち貨幣價値の變化）を知るを目的とするものであるから、比較の基準となるべき或る時期（基準時點）が決定されねばならぬ。基準には固定基準と連鎖基準との別がある。卽ち或る一定時點を共通の基準とし、總てこの時點の物價に對して比較するならば該時點を固定基準といふ。例へば一九〇〇年の物價を基準とし、一九〇一年、一九〇二年等の物價はそれぞれ一九〇〇年のそれに比較せしめるが如きこれである。これに對して、時の經過と共に基準を變更し、常に直前（前年又は前月等）の時點の價格を基準とするものを連鎖基準といふ。連鎖基準は實際には用ひられてゐないから、こゝでは固定基準だけについて說明しよう。

固定基準には一定時點の物價をとる場合と、一定期間の平均物價をとる場合とがある。商工省の「全國卸賣物價指數」は昭和四年十二月の物價を基準とし、反之、ダイヤモンド社の「東京卸賣物價指數」は大正元年八月乃至大正三年七月の二箇年間の平均物價を基準としてゐる。元來、基準は比較の根柢となるものであるから、最も正常的なものでなければならぬ。故に或る一定時點をとつたのでは、何等か偶然の理由によつて異常な物價を示してゐるかも知れぬ。經濟界の比較的靜穩な時期を選べばこの危險は著しく避けうるが、併しその場合にも猶ほ多少の期間を平均した方がより正常な物價を示められる事は、平均の性質上明かである。斯かる基準を廣礎法といふ。

（二）綜合指數の算式　何れにしても基準が決定されゝば、これに對して他の時期の物價を比較すれば良いのであるから、$I_{01}, I_{02}, I_{03} \ldots\ldots I_{0n}$ の形を採るのである（Iは物價指數、oは基準時點の物價、1、2、3、……nは各時點の物價を示す。故にI_{01}は基準時點に對する第一時點の物價を、I_{0n}は基準時點に對する第n時點の物價を示す）。さて各時點の物價とは、既に說明した通り、重要商品の價格の綜合又は平均に外ならぬが、これを指數に改める爲には或る計算を必要とする。大別して總和法と平均法の二つがあり、その各〻が單純法と加重法とに分たれる。

A、總和法　この方法は米國の經濟雜誌社ブラッドストリートのみの採用するもので、極めて特殊の形態を示してゐる。卽ち同社は一〇六商品の各一封度の價格を合計し、これを以て直ちに物價指數となすもので、例へば九弗五十六仙といふが如き形で發表される。指數とは元來比率たるべきものであるから、如上の形態は指數とは言へぬ筈である。且つ斯かる發表形式に於ては基準時は存在しない事になるから、定義的な物價指數とは全く別箇のものと言つてもよからう。しかもこの方法には根本的な缺陷がある。いま問題を簡單にする爲に、商品數を甲乙の二つに限定して考へて見よう。先月の甲商品は一封度十錢、乙商品は五十錢とすれば指數は六十錢である。今月に至つて甲は二割騰

	昨年度	今年度
甲	100	120
乙	100	80
綜合	100	100

貴し乙は二割下落したとすれば、甲は十二錢、乙は四十錢となり、從つて指數は五十二錢となつて先月よりも八錢、即ち約一割三分の下落となる。一封度價格の大小は決して國民經濟的重要さを反映するものではないから、いま假りに甲乙兩商品の重要さに相違なきものとすれば、一方の二割の騰貴は他方の二割の下落によつて相殺されて然るべきである。更にもし甲は乙よりも遙かに重要な商品とすれば、その二割の騰貴は乙の二割の下落によつて相殺されず、物價は事實に於て反つて騰貴した事になる。即ち理論的には總和法は殆ど一顧の價値もないのであるが、而もブラッドストリート指數が今猶ほ米國の代表的物價指數の一つとして尊重されてゐる所以は、一はその歷史の古い爲であり、一は商品選定が實際上極めて巧妙に行はれてゐる爲である。尙ほこの指數は、或る時點を基準とした比率に換算する事は容易であるから、例へば東洋經濟新報社では、大正二年一月を一〇〇とした指數に改算して發表し、他の指數との比較を便ならしめてゐる。

註　ブラッドストリート社は一九三三年ダン商會と合併し、Dun & Bradstreet 指數として發表されるに至つた。

B、平均法　選定された各商品につき、求める月の價格の、基準時の價格に對する比率を求め、これら比率を平均する事によつて綜合指數を求める事をいふ。平均する際に算術平均又は幾何平均が用ひられ、他の平均即ち並數、中位數又は調和平均等は實際に用ひられる例はない。いま前例を以て說明すれば次の如くなる。甲乙二商品に於て、甲は昨年度を一〇〇とすれば今年度は一二〇、乙は昨年度を一〇〇とすれば今年度は八〇であつた。故に昨年度は甲乙共に一〇〇であるから平均は同じく一〇〇である――算術平均でも幾何平均でもこの場合は同じである。蓋し $Ma=\frac{100+100}{2}=100$, $Mg=\sqrt{100\times100}=100$ だからである。然るに今年度は甲は一二〇、乙は八〇であるから、算術平均によれば $(120+80)\div2=100$、幾何平均によれば $\sqrt{120\times80}=98$ となる。實際には商品數は二つや三つではないから、幾何平均を求めるには對數を使用せざるを得ない。

綜合指數の作製に際して算術平均を採るべきか幾何平均を採るべきかは可成り六ケしい問題であるが、理論的には幾何平均が優れてゐる。その主たる理由の一つは基準の變更が容易な事である。簡單に說明して見よう。いまA、B二商品が第一時點には共に五〇圓、第二時點にはAは一〇〇圓、Bは二五圓になつたとする。算術平均によつて第一時點を基準とした綜合指數を求めれば、第一時點は $\frac{100+100}{2}=100$, 第二時點は $\frac{200+50}{2}=125$ となり、後者は前者に較べて二五%の上昇となる。然るに第二時點を基準とすれば、第二時點（A一〇〇圓、B二五圓）の綜合指數は $\frac{100+100}{2}=100$ なるに對し、第一時點（A五〇圓、B五〇圓）のそれは $\frac{50+200}{2}=125$ となり、前の場合とは反對に第一時點の方が二五%だけ高かつた事になる。これでは中途で新たな基準を定め、過去の指數を換算する事は不可能だといふ事になる。强ひて新たな基準を定めるならば、箇々の指數に遡つて計算せねばならないのである。

然るに幾何平均を用ふれば、右の不便は一掃される。即ち同じ例を用ひて說明すれば次の如くなる。第一時點を基準とすれば、第一時點の綜合指數は $\sqrt{100\times100}=100$、第二時點のそれは $\sqrt{200\times50}=100$ であり、次に第二時點を基準とすれば第二時點の綜合指數は $\sqrt{100\times100}=100$、第一時點のそれは $\sqrt{50\times200}=100$ となり、要するに基準を變更しても算術平均に於けるが如き不合理は生じないのである。

註　右例に於ける第一時點を基準とした第一時點の算術平均は一〇〇、第二時點のそれは一二五であり、これに對し第一時點を基準とした第一時點の幾何平均は一〇〇、第二時點のそれは同じく一〇〇である。その何れが正しいであらうか。これは差を重ん

ずるか比を重んずるかによつて答へは自ら違つて來る。卽ち第一時點に於てA、B兩者を賄ふには一〇〇圓で足りたものが第二時點では一二五圓となつたのであるから、指數も二五％上昇するのが正しいといふ見方もあり、これに對して、第二時點に於てはAは二倍に、Bは半分になつたのであるから、兩者は帳消しとなつて指數は依然一〇〇なるが正しいといふ見方もある。前者は差を、後者は比を重んじた結果である。

（三）**加重**　算式決定の第二の問題は「加重」――評量又はウェート（Weight）ともいふ――といふ事である。加重とは、各商品を總て一樣に取扱ふ代りに、その各〻の重要性に比例した取扱ひをなす事である。何故これが必要かは、次の例で判らう。いま商品が米と牛肉の二つだけだと假定し、或る時點に於ける各〻の價格は米は一石三十圓、牛肉は一貫目十圓だつたとする。次の時點は米は三十二圓、牛肉は八圓になつたとすれば、上表の如く米價は一〇六・六と六・六％の騰貴、牛肉價は八〇・〇二〇％の下落となり、平均すれば九三・三となり、六・七％の下落となる。この答へは計算上は正しいが、理窟の上からは極めて怪しいのである。

	第一時點		第二時點	
米	三〇圓	一〇〇％	三二圓	一〇六・六％
牛肉	一〇	一〇〇	八	八〇・〇
		二〇〇		九三・三

我國で一年間に取引される牛肉は五六千萬圓に過ぎないが、反之米は十億圓に達すると推定されてゐる。牛肉一貫目十圓とすれば數量は五六百萬貫であり、米一石三十圓とすれば數量は約三千三百萬石となる。卽ち牛肉價の二圓の下落は一千萬圓そこそこに達するに過ぎないに對し、米價の二圓の騰貴は六千六百萬圓の騰貴を意味する。差引すれば五千六百萬圓程度の騰貴、卽ち貨幣價値の下落が現はれて來なければならぬ筈である。然るに上の計算では全く逆の結果となつてゐる。右は平均法に據つたものであるが、總和法に於ても同種の缺陷が看取される。卽ち第一時點には兩商品の價格合計は 30圓＋10圓＝40圓 であり第二時點のそれは 32圓＋8圓＝40圓 で、全く差がない事になり、貨幣價値下落の跡は毫も認められない。

この不合理は全く各商品の重要さを無視し、總てを一樣に取扱つた結果である。斯くて、この不合理を是正する爲には加重法によつて各商品の重要さを算入する方法を採らねばならぬ。加重法とはこの爲に案出された特殊の方法であるが、併し本來の意味に於ける加重法に對し、加重せざる方法卽ち單純法と雖も實は陰伏性の加重を採用してゐる事を指摘して置きたい。單純法に據つてゐる東洋經濟新報社物價指數の品目を見るに、例へば內地米と朝鮮米とが含まれてゐる。この二つの米は本質的に別に變つたものではなく、從つて價格も甚だ關聯してゐるから、單に米價を見ようとする丈けならばそのどちらかで足りるわけで、特に兩者を擧げて置く必要はあるまい。然るに物價指數に於て米が二品目を占めれば、米價の動きは他の品目の價格の動きに較べて二倍の重要さを附與された事になるのである。故にこの所謂「品目の分割」は實質的には加重法に外ならないのであつて、單純法に據る如何なる物價指數も必ず或る程度まで右の手段を講じてゐるから、若し加重法を廣い意味に解釋すれば、加重法に據らざる物價指數は皆無と言つてよいのである。乍併この原則から行くと品目の分割は殆ど際限なく行はねばならない事になる。例へば米は味噌に對し五十倍の重要さがあるとすれば、五十種の米を擧げねばならず、更に小麥は味噌の二十倍とすれば二十種の小麥を擧げねばならず、斯くては最初に僅か十種か二十種の商品を選擇しても、品目の分割によつて五百にも千にも上つて了ふであらう。この煩を避けて一商品をせいぜい二乃至三品目に分割するに止めれば、不公平は依然として殘らう。本來の加重法とは斯かる品目分割を指すのではない。それには各商品の重要さを豫め或る數字で示し、その數字

をば重さ（評量値、ウェート）として算式中に明示する事を必要とする。素より商品の持つ重要さは時と共に常に變化しつゝあるから、全く正確な評量値は求められるものではなく、實際の問題としてはその近似値で滿足する外はない。この近似的評量値は如何にして求められるか、又如何にして算式に適用されるかを一考して見よう。

加重の方法は算式が總和法であるか、平均法であるかに依つて異り、前者に於ては數量（取引量、消費量等）が、後者に於ては價額（取引價額、消費價額等）が用ひられる。

何故かと言へば總和法でも平均法でも、前例に於て基準價額（米十億圓、牛肉五千萬圓、計十億五千萬圓）に對する比較價額（米十二億圓、牛肉四千萬圓、計十二億四千萬圓）の比率を示し得るやうな加重法を採ればよいのであるから、（一）總和法に於ては單位價格に數量を掛ければ分母及び分子の絶對的價額が求められるし、（二）平均法に於ては單位價格の比率に價額を掛ければ相對的價額が求められ、畢竟この二つは同じものに歸着するのである、加重に際して何時の數量又は價額を用ふるかは極めて六ケしい問題であるが、基準時のそれらが用ひられる場合が最も多いから、その場合の一般の算式を擧げて置かう。價格を p、數量を q、基準時點及び比較時點をそれぞれ小文字の o 及び 1 で示せば例へば p_0 は基準時點の價格、q_1 は比較時點の數量を示す。更に基準時點に對する比較時點の指數を I_{01} で表はせば

（一）加重總和法に於ては

$$I_{01}=\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$$

（二）加重平均法に於ては

$$I_{01}=\frac{\Sigma\left(\frac{p_1}{p_0}\times p_0 q_0\right)}{\Sigma p_0 q_0}=\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$$

となり、結局共じ結果に歸着して了ふ。この式を前例に適用すれば總和法に於ては

$$\frac{32\text{圓}\times 333\text{萬石}+8\text{圓}\times 500\text{萬石}}{30\text{圓}\times 333\text{萬石}+10\text{圓}\times 500\text{萬石}}=\frac{110.6\text{千萬圓}}{105\text{千萬圓}}=105\%$$

平均法に於ては

$$\frac{106.6\%\times 10\text{億圓}+80.0\%\times 5\text{千萬圓}}{10\text{億圓}+5\text{千萬圓}}=\frac{110.6\text{千萬圓}}{105\text{千萬圓}}=105\%$$

となり、5％の上昇（卽ち貨幣價値の下落）を示す。これは從來兩商品の購入に十億五千萬圓で足りたものが差引き五千六百萬圓を增して十一億六百萬圓となつた當然の比率である。既に述べた通り、大部分の物價指數は平均法に據つてゐるから、加重には價額、就中取引總金額が使用される。併し問題となるのは總金額そのものではなく、その相互の比であるから例へば取引金額千萬圓の商品の重要さを一とすれば一億圓の商品の重要さは十とすれば足りるわけで、これらの數字を「重さ」「ウェート」又は「評量値」といふのである。

註　度數分布表から平均を求めるには、算術平均の場合には各階級の中點 m に度數 f を乘じたものを度數の合計で割ればよかつた。卽ち $\frac{m_1f_1+m_2f_2+\cdots\cdots m_nf_n}{f_1+f_2+\cdots\cdots f_n}=\frac{\Sigma(mf)}{\Sigma f}$ であつた。これは要するに度數によつて加重されたものと見てよいのである。又幾何平均の場合には $\sqrt[f_1+f_2+\cdots\cdots f_n]{m_1^{f_1}\times m_2^{f_2}\times\cdots\cdots m_n^{f_n}}=\sqrt[\Sigma f]{\Pi m^f}$ であつたが、これも亦加重幾何平均と同じものである。卽ち基準時點の價額で加重した幾何平均法は $I_{01}=\sqrt[\Sigma p_0 q_0]{\Pi\left(\frac{p_1}{p_0}\right)^{p_0 q_0}}$ であり、對數に書直せば $\log I_{01}=\frac{1}{\Sigma p_0 q_0}\Sigma\left(p_0 q_0\times\log\frac{p_1}{p_0}\right)$ となる。我國のダイヤモンド社指數はこれによつてゐる。

×　　×　　×　　×

算式については幾多重要な理論的問題が横はつてゐるが、私はこれらを平易に説明する事は不可能と考へ、一切割愛した。別の「經濟統計」なる講座に於て、もし紙數に餘裕があつたならば、これに論及したいと思つてゐる。

第四節　物價指數の實例

商品種目及び評量値選定の實例として昭和十一年十二月改正の日本銀行「東京卸賣物價指數」を見よう。同行指數は久しく品目五十六、算式は單純算術平均の方法によつてゐたが、時代の推移と共にその不備が明かとなつたので遂に根本的改正を加へ、品目は一一〇、算式は加重算術平均の方法を採用するに至つた（次表參照）。品目選定に當つては、昭和七年乃至九年の三箇年間の品目別內地生產價額及び同輸移入價額合計の年平均を算出し、それから取引に上らぬ自家用消費推定額を控除したものを以て取引額と見做し、原則としてその額が一千萬圓以上のものを採用したのであるが、將來取引額の引續き增加する見込みのある品目については、現在取引額が一千萬圓以下でも採用した。更に評量値は大體取引額一千萬圓を以て一としてある。即ち例へば內地米の評量値は七十四であるから、取引額は七億四千萬圓と推定されたのであらう。苛性曹達、晒粉等の評量値は〇・五であるから、取引額は未だ一千萬圓に達しない事が判る。これらの調査及び推定は日銀當局者の愼重な手續によつて行はれたものであるが、我國に於ける生產及び消費の統計の未だ甚だ不備な事實を顧れば、これら評量値に過度の信頼を置く事は危險と思はれる。

改正新卸賣物價指數の採用品目及評量値

品　目	評量値	品　目	評量値	品　目	評量値	品　目	評量値
食用農產物	128	生絲	47	×內地材	11	魚肥	3
內地米	74	〇紡績絹絲	4	外國材	4	**工業藥材**	8
〇朝鮮米	17	〇人造絹絲	13	洋灰	8	〇硫酸	1
〇臺灣米	8	×米棉	16	板硝子	3	〇硝酸	1
大麥	3	×印棉	9	瓦	2	〇鹽酸	0.5
裸麥	4	綿絲	48	疊表	1	苛性曹達	2
×內地小麥	8	〇羊毛	7	石材	1	〇曹達灰	1.5
×外國小麥	4	毛絲	10	**金屬類**	83	〇晒粉	0.5
×內地大豆	4	麻	2	〇銑鐵	6	〇氷醋酸	0.5
×滿洲大豆	4	〇麻絲	2	鋼	46	〇グリセリン	0.5
小豆	2	**布帛類**	110	銅（電氣銅）	9	〇鹽酸加里	0.5
其他食料及嗜好品	127	×羽二重內地向	1	鉛　（塊）	2	**其他**	45
小麥粉	13	×　〃　輸出向	2	亞鉛（〃）	2	〇染料	3
×分蜜糖	10	〇縮緬內地向	7	〇錫　（〃）	2	〇塗料	1
×精糖	11	〇　〃　輸出向	4	〇アルミニユーム	3	〇パルプ	3
食鹽	4	絹裏地	2	〇眞鍮	4	×印刷紙	
味噌	2	〇銘仙	5	〇亞鉛鍍板	3	×模造紙	16
醬油	7	〇富士絹	3	〇鉽力	4	×更紙	
牛肉	5	〇人絹縮緬	17	釘	2	〇生ゴム	3
〇豚肉	3	〇人絹平織		**燃料**	45	〇ゴムタイヤー	3
鷄卵	6	〇綾木綿	10	石炭	24	〇絶緣電線	4
鰹節	1	〇粗布	5	重油	4	〇セルロイド素地	2
〇魚類罐詰	6	金巾	17	揮發油	6	〇硬化油	1
〇果實〃	1	〇天竺	4	骸炭	3	〇機械油	2
×大豆油	1	〇綿朱子	3	木炭	8	皮革	5
×種子油	1	白木綿	4	**肥料**	20	燐寸	1
茶	2	毛斯綸	5	大豆粕	5	〇琺瑯鐵器	1
日本酒	25	羅紗	12	菜種粕	1		
〇麥酒	8	〇莫大小	7	硫安	6	**總計**	74
卷煙草	21	〇麻布	2	〇過燐酸	3		
纖維原料品	158	**建築材料**	30	〇石灰窒素	2		

〇は在來の指數に採用せざるもの　×は在來の指數に採用したる品目を分割せるもの

註　品目及び算式を變更すればその瞬間から從前の指數とは別趣の指數が現れる事になり、理論的には新指數を舊指數に接續せしめる事は不可能の筈であるが、日銀では「新方法ニ依リ昭和八年ヲ基準トシテ昭和七年以來最近ニ亘リ作成セル指數ト、同一期間ニ於ケル在來ノ指數ヲ同一基準ニ改算シタルモノト對照シタル結果、兩者ガ昭和九年頃ノ位置ニ多少相違セル所アル以外其ノ動向ヲ同ウシ、殊ニ最近一年間ハ其位置並ニ趨向共ニ殆ド全ク相一致セルコトヲ確メタルヲ以テ此機會ニ於テ兩者ヲ接續シ從來ノ分ト一貫セル指數トシテ發表スルコト、定メタルナリ」と言つて、次の算式によつて依然舊指數に接續せしめて發表してゐる。

$$\text{從來ノ方法ニ依ル昭和11年11月分ノ指數}(203.5)\times\frac{\text{新方法ニヨル昭和11年12月以後ノ指數}}{\text{新方法ニヨル昭和11年11月分指數}(112.9)}=\text{新方法ニヨル昭和11年12月以後ノ指數}\times\frac{203.5}{112.9}$$

$$=\text{新方法ニヨル昭和11年12月以後ノ指數}\times 1.8$$

第五節　生計費指數

物價指數の特殊形態として小賣物價指數及び生計費指數がある。既に述べた通り卸賣物價指數は一般的抽象的貨幣價値の尺度であるが、これに對して吾人の日常使用する貨幣は主として小賣商品に支出されるから、斯かる支出項目から具體的貨幣價値を算出する事が出來、またその必要がある。小賣物價指數とは普通に賣買される主要商品の小賣價格の變動を示すを目的とするもので、例へば日銀東京小賣物價指數の如きものが有名である。併し更に一歩を進めて考へれば、小賣商品だけが吾人の日常の支出項目ではない。家賃とか召使の給料とか、又は租税・授業料とか、所謂普通の小賣商品以外の支出項目が尠くない。故に實際の具體的貨幣價値を測定するには、これら日常の用途を總て考慮に入れなければならぬ。斯かる用途に支出される毎月の金額を生計費と名づけ、右の目的に應ずる爲に生計費指數なるものが作製されるのである。

然るにこの指數の作製には著しい困難が伴ふ。それは、家庭毎に收入・職業・社會的地位　家族の員數や年齡・嗜好・習慣等が相違するから生活樣式も亦當然相違し、支出の項目も割合も千差萬別であり、從つて生計費指數作製に際してどの家庭にも通用する品目や評量値を決定するが如きは全く不可能と言つてよいのである。そしてこの生活樣式の差は富裕階級の間で特に甚だしい。金錢が自由になるから趣味や贅澤に湯水のやうに浪費する人もあらうし、又は反對に蓄れば蓄るほど穢くなつて爪に火をとぼすやうな連中も出て來るわけで、彼等の支出狀態は全く箇々別々である。これが貧困階級になると、趣味や贅澤や貯蓄なぞを顧る暇はなく、收入の大部分は否應なく生活必需品に振り當てられねばならぬから、自ら或る共通的生活樣式をとることになり、支出狀態に一定の基準を發見する事が出來る。

理由の第二は、作製の目的から說明する事が出來る。自分の持つ貨幣價値の變動は誰に取つても氣になるものであるが、それが最も深刻な關係を持つのは、言ふ迄もなく貧困階級の間に於てゞある。餘裕ある人々ならば假令貨幣價値が下つても從來の無用の支出を多少切りつめれば依然安樂な暮しが出來ようが、貧民に取つては格別切りつめて差支へないやうな項目は最初から無いのであるから、忽ち生活を脅かされ、貧困は赤貧に、赤貧は餓死に轉落せねばならぬ。これら不幸の境遇を一掃する事こそ國家理想の中心たるべきであり、この爲にこそ各種の社會政策が立案施行されてゐるのである。而も次第に加はる人口の壓力や避け難い戰時體制の强化なぞは大衆の貧困化を促進する傾きがある。然らば彼等の生活狀態の變化を正確に測定する事は、特に非常時局下に於て特別の重要性を帶びるのである。

右に述べた二つの理由から、生計費指數は常に貧困階級の生計費の變動を示すやうに作製される。それには先づ彼等が收入を如何なる項目に如何なる割合に支出するかを決定せねばならぬ。これは机の上で出來るものではなく、實

本邦生計費指數品種別全國標準評量值（內閣統計局計算）

品種別	評量値
總計	一、〇〇〇・〇〇
飲食物	四八〇・三〇
米麥類	一三五・六六
白米	一三〇・三三
麥	三・〇六
食麺麭	二・二八
魚介類	二四・三三
鰯	一・七八
鯖	三・三三
鰈	二・〇二
鯛	七・五八
鹽鮭	二・〇一
干鰈	〇・三三
干鰯	〇・九二
煮干	一・五九
鯡	一・六一
鰊	〇・六九
烏賊	一・八五
目刺	〇・六三
•肉類	一六・〇九
牛肉	一一・九七
豚肉	一・五七
鷄肉	二・五五
•鷄卵	九・九四
•牛乳	九・五一
豆及蔬菜類	二五・四三
大豆	〇・八四
小豆	一・一七
菜豆	一・一〇
玉葱	一・〇四
馬鈴薯	二・五〇
大根	四・八三
漬菜	〇・六五
甘藷	二・四二
人參	〇・八〇
牛蒡	一・六六
里芋	二・三四
菠薐草	〇・七七
茄子	二・四五
胡瓜	一・九六
玉菜	〇・九〇
乾物類	六・四九
海苔	二・〇二
昆布	三・四五
干瓢	〇・四〇
麩	〇・六二
豆腐・佃煮・煮物・漬物類	一八・三三
豆腐	七・一一
油揚	二・八七
蒟蒻	二・〇一
梅干	〇・四六
澤庵	一・五五
福神漬	〇・三八
奈良漬	〇・九五
竹輪	一・六四
豆腐カラ	〇・一三
煮豆	一・二三
調味料	四六・六八
味噌	七・〇六
醤油	二〇・二一
鹽	一・四三
砂糖	一三・一二
鰹節	四・八七
饂飩	二・三九
嗜好品	九五・四六
酒類	四〇・〇八
淸酒	三四・四九
麥酒	五・五九
•煙草	二一・五二
菓子果物類	二六・五二
•煎餅	五・一五
大福	一・七六
柿	一・二三
•バナナ	二・二八
餡パン	一・〇六
羊羹	〇・九五
•ビスケット	二・六三
飴菓子	二・四四
キャラメル	一・四三
燒芋	〇・七七
林檎	二・三五
蜜柑	三・二四
梨	一・二三
飲料其他	七・三四
サイダー	一・六一
•茶	三・八六
氷	一・八七
住宅	二二〇・七〇
家賃(又ハ間代)	二一二・八八
住宅修繕	四・四三
疊表	一・六六
疊裏返	〇・八六
疊上敷	〇・八三
障子紙	一・〇八
水道	三・三九
光熱	六八・八八
木炭	三〇・三七
薪	一〇・九三
石炭	二・二六
瓦斯	六・七〇
電燈	一六・〇二
炭團	二・二九
燐寸	〇・三一
被服	一三四・七九
衣服類	八〇・五九
甲斐絹	〇・四九
銘仙	一一・一六
木ネル	三・五三
モスリン	一〇・六七
毛絲	二・〇九
セル	三・三四
•洋服	二・五七
白木綿	一・九三
木綿織絣	一・七四
木綿縞	一・〇〇
木綿染絣	一・四五
金巾裏地	〇・四九
綿縫絲	〇・五九
•綿	六・二三
•浴衣地	三・一八
子供洋服	一〇・〇七
雨合羽	〇・三一
綿打直し	〇・七五
身の廻り品	五四・二〇
莫大小シャツ	七・四五
足袋	五・七一
•下駄	一〇・八八
雨傘	一・四八
洋傘	一・九四
•靴	一一・四〇
靴下	二・七九
帽子	・六三
浴巾	〇・九六
ワイシャツ	一・三三
カラー	〇・三〇
猿又	一・〇六
半襟	一・〇四
草履	一・九〇
靴墨	〇・二四
其他	九五・三三
保健衛生	四三・一三
石鹸	六・九二
處方感冒藥	〇・一四
處方胃腸藥	〇・〇七
感冒賣藥	〇・九五
胃腸賣藥	〇・六七
クリーニング及洗濯	三・三〇
入浴料	一七・五四
齒磨粉	〇・四九
理髪	九・二一
齒刷子	〇・四〇
膏藥	〇・三〇
口中藥	〇・二三
白粉	〇・九八
洗粉	〇・三〇
髪油	一・二三
ポマード	〇・四〇
育兒(玩具)	三・六二
教育	〇・九二
鉛筆	〇・三八
インク	〇・一八
墨	〇・一一
筆	〇・二五
交通(電車汽車)	二四・二一
文房具	一・一九
半紙	〇・四五
ペン先	〇・〇八
雜記帳	〇・五六
封筒	〇・一〇
修養娯樂	二〇・一五
新聞	一三・五〇
芝居映畫見物	四・一六
雜誌	二・四九
家具什器及設備	二・一一
茶碗	一・〇〇
箒	〇・五三
桶	〇・二三
バケツ	〇・三五

際について調査する外はない。この調査を生計調査といふ。生計調査を行ふ爲には若干の收入階級と、各階級の標準

家族とを定め、その家族をして毎月の收支狀態を詳しく家計簿に記入せしめるのである。收入階級を分つのは、同じく貧困階級といつても各種の段階があり、例へば月收三十圓の家庭と五十圓の家庭とでは生計狀態に格段の差があるからである。また標準家族を定めるのは、同じ收入でも家族構成狀態が餘りに異常であつては生活狀態も自ら異常となつて、從つてその屬する收入階級の一般狀態を反映しないからで、例へば獨身者とか七八人も子供のある人とかは標準として不適當なわけである。斯くて一般には夫婦と子供二三人の家族を標準家族として選定するのである。

我國に於ける代表的生計調査は內閣統計局の行ふところで、大別して給料生活者、勞働者及び農業者とし、五十圓未滿、六十圓未滿……等の收入階級につき略〻一世帶四人乃至六人位ゐの標準家族につき一箇月平均の收入及び支出の詳細を調査するのである。この調査は絕えず繼續され、每年その結果が發表されてゐるが、併し生計調査を生計費指數の基礎とする爲には一應品目及び評量值を確定せねばならぬ。これは恰も卸賣物價指數の作製に於て、各重要商品の重要さ（取引額）は絕えず變化してはゐるが、一應或る時期について品目及び評量值を確定する必要あると同樣である。斯くて內閣統計局は大正十五年九月から翌年八月迄の給料生活者及び勞働者の生計調査に基き、全國及び地方別の標準商品並びに評量值を決定した。理想的には各收入階級每にこれを定め、階級別生計費指數を作るべきであるが、實際問題としては不可能であるから、平均的な一種類に止めて了つた。併し內容は頗る詳細を極めて、その全國標準の分を見るに、別表の如く品目一五三につき各〻支出の十萬分の一までの評量值が決定されたのである。朝日新聞社は略〻この評量值に基き昭和六年十月以降每月全國生計費指數を發表してゐる。基準は大正三年である。また昭和十三年七月から內閣統計局で獨自の生計費指數を發表し始めた。

第八章　相關々係

第一節　相關々係の意義

今までに述べて來たものは何れも、與へられた或る一統計系列を如何に解析するかといふ事であつた。卽ち度數分布については分布形態の解析（平均值、散布度、對稱度なぞの測定）を、時系列については變動形態の分析（長期傾向、季節變動、循環運動、不規則變動の算出又は除去、並びに指數の原理）を問題とした。故に讀者は與へられた各種統計系列を箇々に處理する方法については既に一應の理解は持たれた事と信ずる。

併し若し統計的解析が箇々の統計系列についてしか行へないとするならば、科學的研究法としての統計的方法なるものは、大した威力を誇る事は出來ないであらう。蓋し箇々の事象は、それが自然的事象なると社會的事象なるとを問はず、事實に於ては相互に遊離して獨立に存在する事はあり得ず、互に相關聯し、互に原因となり結果となつて複雜な關係に立つてゐるからである。總て吾人が或る事象を科學的に研究する窮局の目的は、その事象は如何にして起るか、また如何なる結果を生むかを明かならしめるに在る。然るに一事象は幾多事象の作用から起り、また幾多事象に影響を與へるものであるから、もし單に該事象だけを研究對象としたのでは、その構成や變化の狀態は判つても、その據つて來る所以もその及ぼす影響も之を知るに由が無いであらう。斯くては、吾人の目的とは頗る距離の遠いも

のとなり、その效果も推して知る可きである。

斯くて吾人は更に一歩を進めて異る統計系列を同時に取扱ひ、以て相互の關聯を明かならしめねばならぬ。以下に說明する相關々係の理論は實にこの問題を解かんとするものである。併し翻つて思ふに吾人の科學的努力の窮局の目的は、單に現象間の相互關係を明かにする事に在るのではない。その目的は常に現象を理解する事であるが、理解とは現象のよつて來る原因とその及ぼす作用を知る事を言ふ。この原因と結果の一聯の關係、即ち因果關係こそ吾人の知的追求の終點である。併し一定の原因があれば必ず一定の結果が生ずると言ひ得るのは、それが科學的法則性に律せられてゐる場合に限る。林檎が枝を離れゝば地に落ちると斷言出來るのはそれが引力の法則に律せられてゐるからであり、人が首を刎ねられゝば死ぬと斷言出來るのはそれが生物學の法則に律せられてゐるからである。故に知的追求の終點は、別の言葉で言へば現象を律する法則の探究に外ならない。

こゝで私は冒頭に述べた自然法則と社會法則との相違に立ち戻らねばならぬ。自然現象は普遍妥當の自然法則によつて律せられてゐるに反し、社會現象は畢竟自由意思によつて左右される人間行爲の集積に外ならず、即ち社會現象にはせいぜいのところ或る傾向が認められるに過ぎないのであつて、この傾向をば一般に社會法則と名づけるのである（二―三頁參照）。斯く社會現象に嚴格な法則が缺如してゐる當然の結果として、社會現象の間には一義的な因果關係は認識し得ないといふ事になる。これを敷衍すれば社會科學は最も重要な點で科學性を失つてゐるといふ事になる。併し乍ら私はケトレーを論ずるに當り、而も大數觀察の結果によればこの社會的傾向なるものが概して驚くべき正規性を帶びてゐること、從つて普遍妥當の自然法則と著しく接近してゐる事を認め得る旨を明かにした。然らば社會科學の領域に於ても、この正規性さへ發見し得るならば、一定の原因があれば一定の蓋然的結果が生ずる事は安じて言ひうる事とならう。換言すれば社會科學の領域に於ても因果關係の探究は必ずしも不可能視さるべきものではないと言ひ得よう。

そこで問題は因果關係を發生せしめるところの法則乃至正規性は如何にして認識されるかといふ事になる。いま私は枝を離れた林檎が地に落ちるのは引力法則があるからだと言つたが、併し始めてニュートンが引力法則を發見したのは彼が林檎が地に落ちるのを見てゞある。換言すれば現象から法則が歸結されたのである。然らば社會的因果關係を律する社會的正規性も亦、現象から歸結されねばならぬ事は明かであらう。具體的に言ふならば、もし二箇の社會現象の間に殆ど普遍的に或る關聯が立證されたならば、そこで始めて兩者は或る法則性に律せられてゐる、或ひは兩者の間には因果關係が存在すると認められるのである。例へばもし物價の上下と賃銀の上下とが必ず並進する事實が遍く立證されるならば、この兩者は或る社會的法則の制約下に在ること、從つて兩者の間に因果關係の存在する事が歸結されるのである。

然るに既に度々述べた通り、社會現象間に全く普遍的な關聯は認められない。前例の物價と賃銀との關係を見ても、物價が上れば多くの場合賃銀も亦上るけれど、時には下る事もある。そこで吾人は先づ第一に、現象間に果して幾何の關聯があるかを決定する必要に迫られるのである。斯かる關聯を指して相關々係（Correlation）といふ。この關聯が判らぬ限り、現象を律する正規性も、從つて因果關係も判る筈はない。即ち相關々係は社會科學に於ける因果關係探究の不可缺の道程なのである。惟ふに因果關係とは相關々係の一特殊形態に外ならぬ。現象間に相互的關聯があり、

のとなり、その效果も推して知る可きである。

斯くて吾人は更に一歩を進めて異る統計系列を同時に取扱ひ、以て相互の關聯を明かならしめねばならぬ。以下に說明する相關々係の理論は實にこの問題を解かんとするものである。併し翻つて思ふに吾人の科學的努力の窮局の目的は、單に現象間の相互關係を明かにする事に在るのではない。その目的は常に現象を理解する事であるが、理解とは現象のよつて來る原因とその及ぼす作用を知る事を言ふ。この原因と結果の一聯の關係、卽ち因果關係こそ吾人の知的追求の終點である。併し一定の原因があれば必ず一定の結果が生ずると言ひ得るのは、それが科學的法則性に律せられてゐる場合に限る。林檎が枝を離れゝば地に落ちると斷言出來るのはそれが引力の法則に律せられてゐるからであり、人が首を刎ねられゝば死ぬと斷言出來るのはそれが生物學の法則に律せられてゐるからである。故に知的追求の終點は、別の言葉で言へば現象を律する法則の探究に外ならない。

こゝで私は冒頭に述べた自然法則と社會法則との相違に立ち戻らねばならぬ。自然現象は普遍妥當の自然法則によつて律せられてゐるに反し、社會現象は畢竟自由意思によつて左右される人間行爲の集積に外ならず、卽ち社會現象にはせいぜいのところ或る傾向が認められるに過ぎないのであつて、この傾向をば一般に社會法則と名づけるのである（二—三頁參照）。斯く社會現象に嚴格な法則が缺如してゐる當然の結果として、社會現象の間には一義的な因果關係は認識し得ないといふ事になる。これを敷衍すれば社會科學は最も重要な點で科學性を失つてゐるといふ事になる。併し乍ら私はケトレーを論ずるに當り、而も大數觀察の結果によればこの社會的傾向なるものが概して驚くべき正規性を帶びてゐること、從つて普遍妥當の自然法則と著しく接近してゐる事を認め得る旨を明かにした。然らば社會科學の領域に於ても、この正規性さへ發見し得るならば、一定の原因があれば一定の蓋然的結果が生ずる事は安じて言ひうる事とならう。換言すれば社會科學の領域に於ても因果關係の探究は必ずしも不可能視さるべきものではないと言ひ得よう。

そこで問題は因果關係を發生せしめるところの法則乃至正規性は如何にして認識されるかといふ事になる。いま私は枝を離れた林檎が地に落ちるのは引力法則があるからだと言つたが、併し始めてニュートンが引力法則を發見したのは彼が林檎が地に落ちるのを見てゞある。換言すれば現象から法則が歸結されたのである。然らば社會的因果關係を律する社會的正規性も亦、現象から歸結されねばならぬ事は明かであらう。具體的に言ふならば、もし二箇の社會現象の間に殆ど普遍的に或る關聯が立證されたならば、そこで始めて兩者は或る法則性に律せられてゐる、或ひは兩者の間には因果關係が存在すると認められるのである。例へばもし物價の上下と賃銀の上下とが必ず並進する事實が遍く立證されるならば、この兩者は或る社會的法則の制約下に在ること、從つて兩者の間に因果關係の存在する事が歸結されるのである。

然るに既に度々述べた通り、社會現象間に全く普遍的な關聯は認められない。前例の物價と賃銀との關係を見ても、物價が上れば多くの場合賃銀も亦上るけれど、時には下る事もある。そこで吾人は先づ第一に、現象間に果して幾何の關聯があるかを決定する必要に迫られるのである。斯かる關聯を指して相關々係(Correlation)といふ。この關聯が判らぬ限り、現象を律する正規性も、從つて因果關係も判る筈はない。卽ち相關々係は社會科學に於ける因果關係探究の不可缺の道程なのである。惟ふに因果關係とは相關々係の一特殊形態に外ならぬ。現象間に相互的關聯があり、

更にその上に右の關聯に於て一が原因で他が結果なる事が確定された場合に始めて因果關係を言ひ得るのである。後に説く通り、統計學は相關々係を決定する事は出來るが、更に進んで因果關係までを決定する能力はない。一言にして言へば統計學は社會科學に對し因果關係を決定する素材を提供するに過ぎない。何となれば統計は元來量的經驗の記述であるから、常に場所と時間とに制約され、從つて因果關係といふやうな一般的關係は統計自體からは生じて來る筈がないのである。では何がこれを決定し得るかといへば、吾人の組織的思考、即ち理論に外ならぬ。この事は後段更に説明するであらう。

斯くて統計學の課題は量的現象の間に相關々係の存在するか否か、もし存するならばその強度如何といふ事である。以下相關圖、回歸線及び相關係數の三項目に分つて、この手續を明かにしよう。

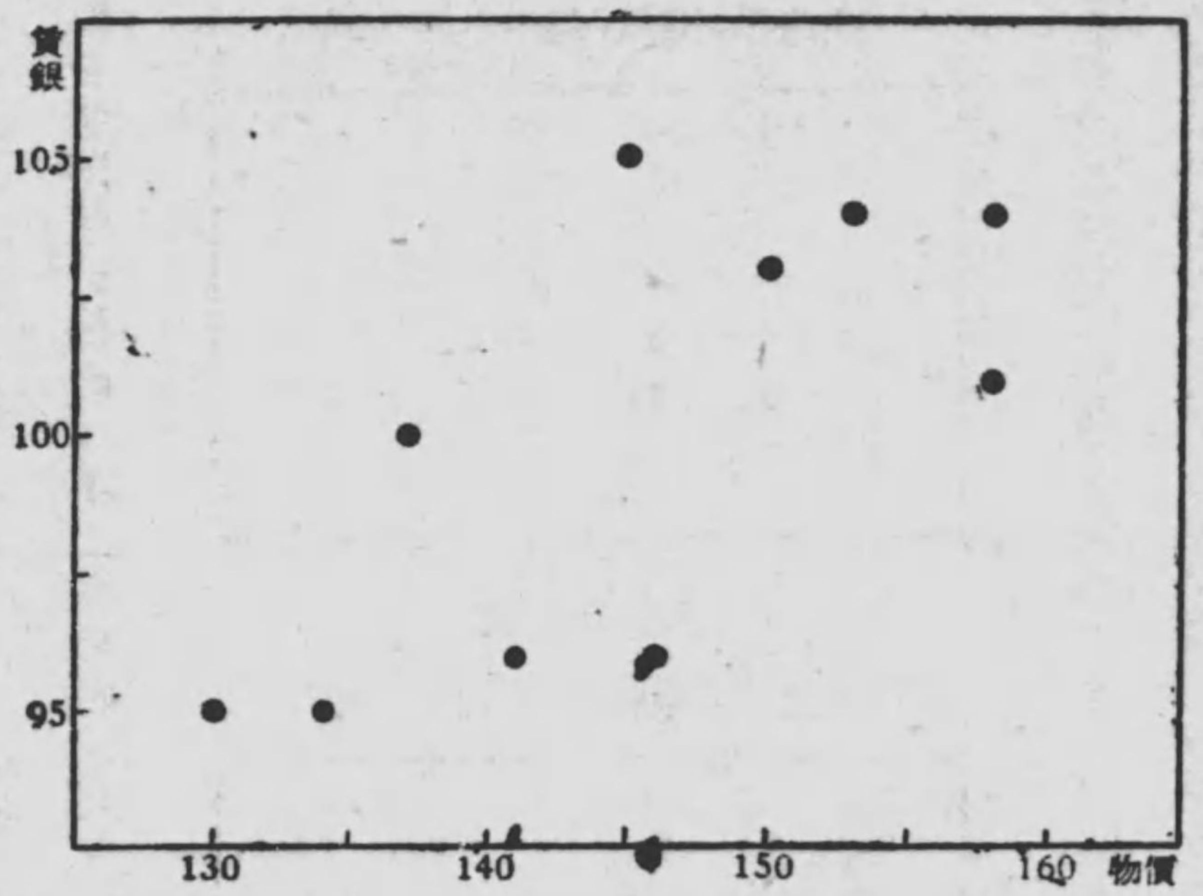
第二節　相關圖

與へられた二箇の量的現象、即ち二箇の統計系列間に相關々係があるか否かは、相關圖（Scatter Diagram）なる特殊の圖表を作つて見れば容易に見當がつく。相關圖とは二系列の一方を縱軸に、他の一方を横軸にとり、二系列の關係をば座標點によつて示した圖表をいふ。茲に最も簡單な例によつて説明しよう。上表は十年間の物價と賃銀とを示す假設例であるが、この表の物價の欄と賃銀の欄とを對比するに、兩者の間に一義的な關係即ち嚴格な數學的函數は認められないが、

年次	物價 X	賃銀 Y
1	158	101
2	146	96
3	141	96
4	134	95
5	130	95
6	137	100
7	150	103
8	158	104
9	153	104
10	145	105

常に必ずしも然りとは言へない。右表に於ても第十年目に於ては前年よりも物價は下り乍ら賃銀は却つて上つて居り、また第三、五及び九年目に於ては前年よりも物價は下つて居り乍ら、賃銀は少しも動いてゐない。併しその他の年について見れば、物價の上つた年には賃銀も上り、物價の下つた年には賃銀も亦下り、所謂順相關の存在する事を示してゐる。與へられた材料が斯くの如く僅かであれば、單に表を一覽する事によつて相關々係の存否を略〻推察し得るであらうが、材料が豐富となれば目測では中々見當がつかない。斯かる場合には上圖の如き相關圖を必要とする。二直線を直角に交らしめ、X軸の目盛は物價を、Y軸の目盛は賃銀を示すものとする。第一年目は物價は一五八、賃銀は一〇一であるから、$x=158$, $y=101$ の座標點によつて第一年目の物價と賃銀との關係を示す事が出來る。第二年目のそれは $x=146$, $y=96$ の座標點、第三年目のそれは $x=141$, $y=96$ の座標點で示されるであらうし、以下同樣にして、與へられた十年間の各年の狀態をそれぞれ座標點として示す事が出來る。

第一圖

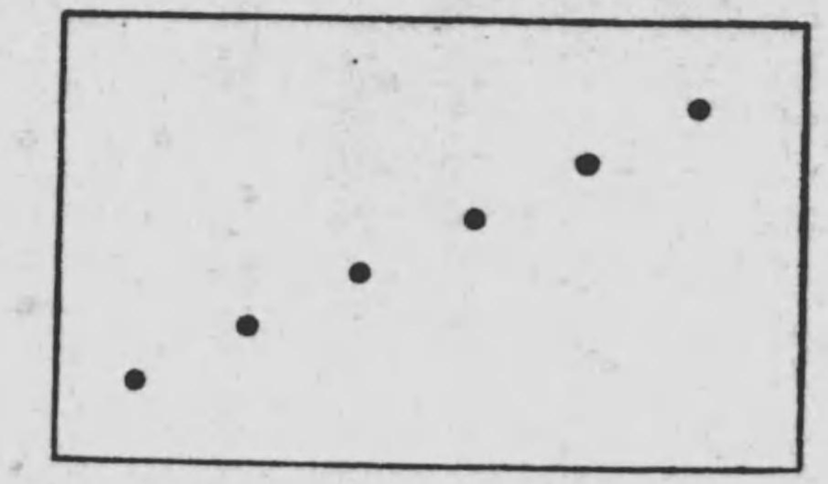


さて若し物價と賃銀との間に最も完全な相關々係（即ち $y=f(x)$ なる數學的函數關係）があるものとすれば、これら各座標點は一直線上に位ひする筈である（複雜な函數關係ならば一曲線上に位ひする）。例へば兩者間に $y=25+0.5x$ の關係があるとすれば、物價が一三〇ならば賃銀

は $25+0.5\times130$ 即ち九〇であり、物價が一五〇ならば賃銀は $25\times0.5\times150$ 即ち一〇〇となるから、斯かる一聯の座標點を相關圖に描けば第二圖の如く總て一直線上に位ひするのである。反之、もし兩者間に何の相關々係も無ければ、

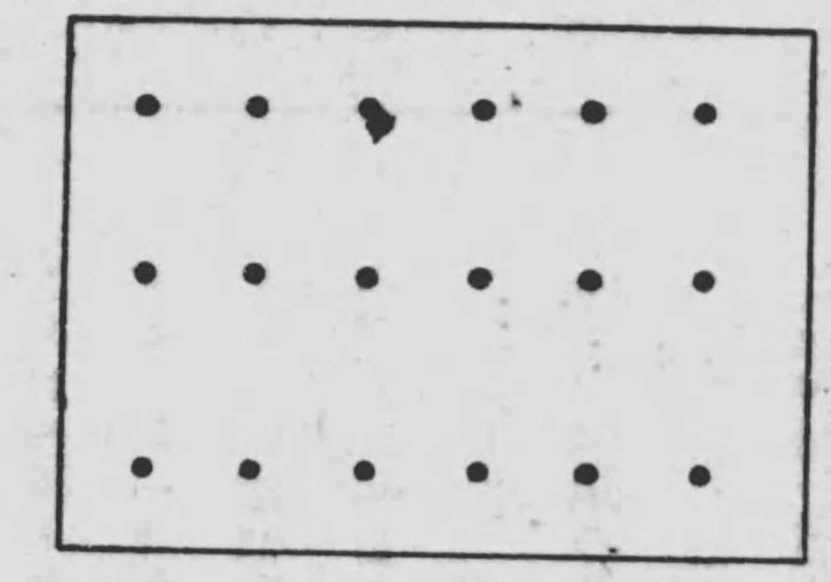
第二圖

完全な相關々係のある場合



第三圖

完全な相關々係のない場合



換言すれば一定の物價の下に高い賃銀も低い賃銀もあるならば、座標點は第三圖の如く平面上に散在して了ふ。

惟ふに二系列間に成り立つ相關々係は、完全な函數關係である場合は、少くとも統計的研究に於ては先づ皆無と言つてよい。併し多少とも相關々係があれば座標點が全體に散在する事もない。即ち實際には相關圖上の座標點は右兩極端の中間的な形を示し、一直線(又は一曲線)上に位ひする事もなく、又全く全面に散在する事もないのである。そして相關々係が大なれば大なるほど各座標點は一直線(又は一曲線)に集中し、これに反して相關々係が小さければ小さいほど集中傾向が少くなつて散在的となるから、要するに相關圖上の各座標點の狀態を見れば、果して兩系列間にどの程度の相關々係があるかを略〻推測する事が出來るのである。

第三節 回歸線

相關々係の有無は上記の如き相關圖の上から判斷されるが、いま二系列間に相關々係の存在する事が判つたとして、更に一變數に對應する他の一變數の最も確からしい數値、即ち理論値を知り得ないものであらうか。然るに例へば兒童の體格檢査に於て男女各歳につき身長は何糎、體重は何瓩といふやうな所謂「標準體格」があるが、これは要するに各年齢に對應する最も確からしい身長又は體重に外ならない。して見れば上記の物價と賃銀についても、物價の或る値に對應する最も確からしい賃銀が無ければならぬ。然らば斯かる理論値は如何にして求められるかと言ふに、時系列解析に於て取扱つた長期傾向値の算出と全く同じ手續を施せばよいのである。最も合理的な長期傾向値は數學線の當嵌めによつて求められる事は既に説明したが、その場合には、時なる自變數とそれに對應する從屬變數とによつて決定される各座標點を最もよく代表する直線又は曲線を引けばよいのであつて、これが爲に最小自乘法を適用したのであつた。上記相關圖に示される座標點はX軸が時の經過を示すものでない點で、時系列の示す座標點とは異るが、二變數間の關係を示す點では毫も異るものではない。長期傾向線は時に對應する變數の最も確からしい値を示す線であるから、全く同樣にして例へば物價に對する賃銀の最も確からしい値を決定する事が出來る。即ち上記の相關圖上の各座標點につき最小自乘法を適用して直線又は曲線を引けばよいのであつて、唯だ斯くして得られた線は傾向線とは呼ばれず、一般に回歸線(又は退行線、Line of Regression)と呼ばれるのである。

傾向線算出に當つて、直線の場合には、與へられた時系列につき $y=a+bx$ なる x に關する一次方程式を求めればよかつた。それに必要な公式は

$$\begin{cases}\Sigma(Y)=Na+b\Sigma(X)\\ \Sigma(XY)=a\Sigma(X)+b\Sigma(X^2)\end{cases}$$

であつたが、この場合のXは必ず時を示す自變數であつた。然るに上記の相關圖ではX軸は物價を、Y軸は賃銀を示すものとしたから、次の如き計算となる。

	X	Y	XY	X^2
1	158	101	15958	24964
2	146	96	14016	21316
3	141	96	13536	19881
4	134	95	12730	17956
5	130	95	12350	16900
6	137	100	13700	18769
7	150	103	15450	22500
8	158	104	16432	24964
9	153	104	15912	23409
10	145	105	15225	21025
計	1452	999	145309	211684

この結果を右公式に當嵌めれば

$$\begin{cases} 999=10a+1452b \\ 145309=1452a+211684b \end{cases}$$

これを解いて

$$\begin{cases} a=56.7 \\ b=0.297 \end{cases}$$

となるから、求める回歸線の方程式は

$$y=56.7+0.297x$$

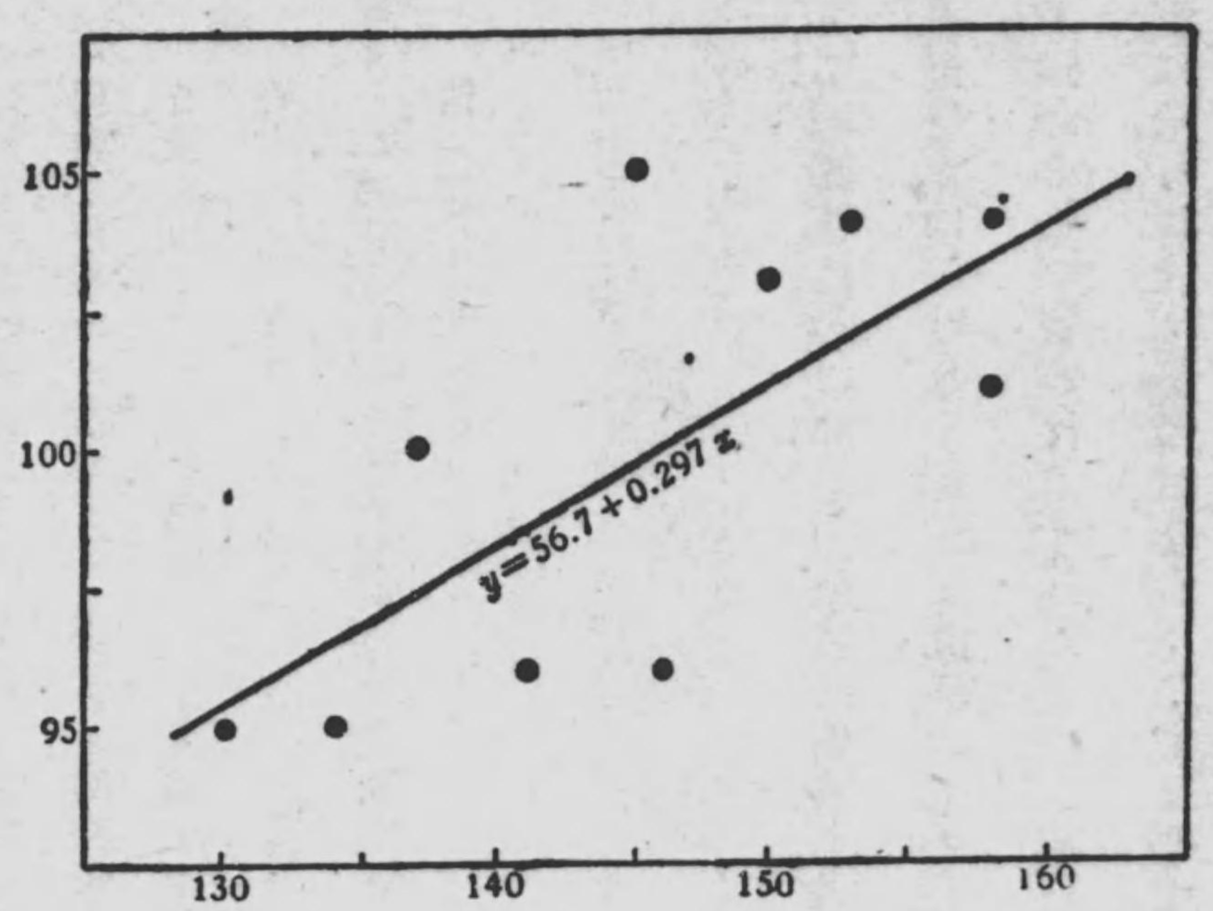
となる。斯くて或る物價に對應する最も確からしい賃銀は、この方程式のxに該物價を代入すれば求められるのであつて、例へば物價が一三〇のときは、賃銀の理論値 yは

$$y=56.7+0.297\times130=95.3$$

であり、實際値たる九五と比較すれば〇・三だけの偏差がある。同樣の計算をXの各値について行つた結果は、次節の表（一六七頁）の「理論値」の欄を見られたい。

右の計算では物價に對應する賃銀の理論値を求めたのであるが、相關圖に於ては何れの變數をX軸に取るも自由であるから、もし賃銀をX軸に取れば同樣の計算によつて賃銀に對應する物價の理論値を求める事が出來る。その計算には右に於けるXとYとを入れ替へればよいから、次の如くなる。

第四圖



	X（賃銀）	Y（物價）	XY	X^2
1	101	158	15958	10201
2	96	146	14016	9216
3	96	141	13536	9216
4	95	134	12730	9025
5	95	130	12350	9025
6	100	137	13700	10000
7	103	150	15450	10609
8	104	158	16432	10816
9	104	153	15912	10816
10	105	145	15225	11025
計	999	1452	145309	99949

$$\begin{cases} 1452=10a+999b \\ 145309=999a+99949b \end{cases}$$

これを解いて　$a=84.26$　$b=0.61$　となるから、求める回歸線の方程式は

$$y=84.26+0.61x$$

となる。これは前述の如く、賃銀に對する物價の理論値を示すもので、例へば、賃銀が一〇〇のときには、物價の理論値は　$84.26+0.61\times100=145.26$　であり、賃銀が一〇五のときには、後者は　$84.26+0.61\times105=148.31$　となる

のである。

註1　二系列間に回歸線を描く場合に、何れをX軸に取る可きかは、時に極めて困難な問題となる。理論的に言ふならば、X軸に取られた系列には誤謬がなく、Y軸に取られた系列には誤謬があるものと假定されねばならぬ。前例で言へば、もしX軸に物價を取れば、該物價統計には誤差がないものと假定されたのである。蓋しX軸は基準となるものであるから、座標點が總て回歸線上に位ひしないのは、Y軸に示される賃銀の側に於ける誤差に基因すると認めざるを得ないからである。故にもし兩系列に同じ程度の誤差が介入してゐると信ぜられる場合には、その何れを基準とするも不可であつて、斯かる場合には、寧ろ各座標點そのものを基準とする特殊の回歸線を描かねばならぬ。これを相互回歸線 (Line of Mutual Regression) といふ。これを描くには、兩軸を無視し、單に各座標點からの距離の平方の總和が最小となるが如き線を決定するのである。この問題については H. Schultz—Statistical Laws of Demand and Supply, 1928 を參照されたい。

註2　二系列間に、一方が增加すれば他方はそれに伴つて減少する關係があれば、回歸線は下降的となる。斯かる相關々係を逆相關 (Inverse Correlation) といふ。

第四節　標準誤差

さて回歸線は右の如く一系列に對應する他の系列の最も確からしい値を示すものであるから、その論理的性質は恰も度數分布に於ける平均値と同樣のものである。某工場の女工賃銀が平均八十錢だといふ意味は、該工場の各女工の最も確からしい賃銀は八十錢だといふ事に外ならぬ。然るに各女工の實際の賃銀は必ずしも八十錢ではなく、事實は一部はそれ以上、一部はそれ以下であつて、要するに八十錢なる平均値は理論的抽象的數値に過ぎない。この平均値が幾許の程度まで事實と合致するか、換言すれば幾許の程度まで信賴し得るかは、散布度の大小によつて決定されること、そしてその爲には標準偏差(σ)の用ひられること等については既に說明した。さて回歸線は平面上に散在する各座標點を代表する「最も確からしい」値を示すものであるから、該線は時には各座標點を能く代表することもあるし、時には充分に代表しないこともあらう。故に幾許の程度までこれに信賴を置き得るかは、平均の問題に於ける標準偏差と同樣の手段によつて決定されねばならぬ。

然るに標準偏差は平均（算術平均）と各項との差(d)の平方の總和をば項の數(N)で除したものゝ平方根、即ち

$$\sigma=\sqrt{\frac{\Sigma(d^2)}{N}}$$

であつた。この方法は回歸線の場合にもそのまゝ適用出來るのであつて、異る點は單に「平均」の代りに「理論値」を取ればよいのである。そして斯くて得た測定値は、標準偏差と區別する爲に標準誤差 (Standard Error of Estimate) と呼ばれ、Sの記號で示される。いまこれを前例について求めれば上の如くである。

年次	X	Y	理論値	d	d^2
1	158	101	103.6	−2.6	6.7
2	146	96	100.0	−4.0	16.0
3	141	96	98.5	−2.5	6.2
4	134	95	94.9	+0.1	0.01
5	130	95	95.3	−0.3	0.09
6	137	100	97.4	+2.6	6.0
7	150	103	101.2	+2.8	7.8
8	158	104	103.6	+0.4	0.16
9	153	104	102.1	+1.9	3.6
10	145	105	99.7	+5.3	28.0
					75.3

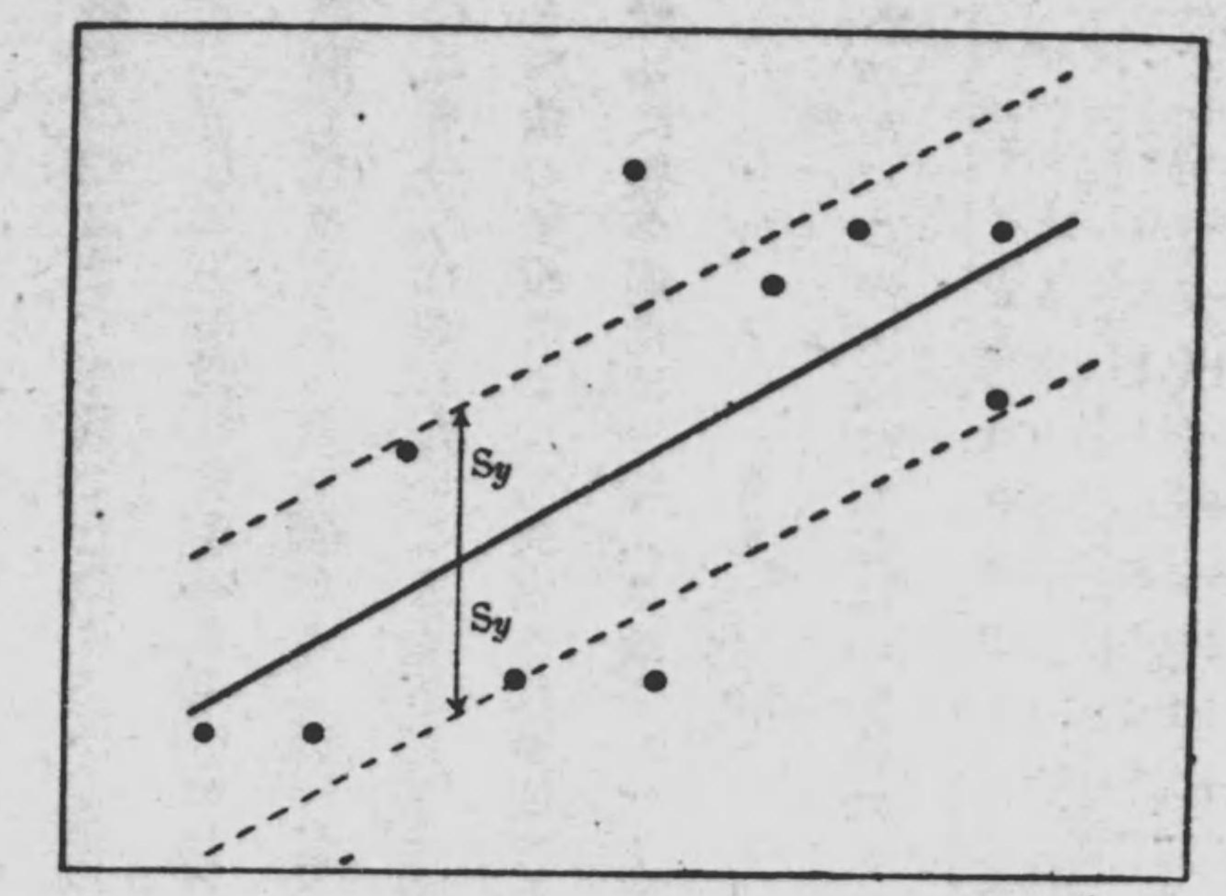
$$S_y=\sqrt{\frac{\Sigma(d^2)}{N}}$$

$$=\sqrt{\frac{75.3}{10}}=2.7$$

S_yと記したのは、上表は物價Xに對する賃銀Yの理論値を問題とするものであるから、標準誤差も亦Xに對するYの標準誤差を意味するからである。さてこの標準誤差は如何に解釋さるべきか。

正常度數分布に於て、平均を中心としてその上下に標準偏差に等しい値を取れば、その限界內に全度數の約七割（詳しくは六八・二六％）が含まれ、標準偏差の三倍を取れば殆ど全部（詳しくは九九・七％）が含まれる事は既に説いた。例へば兒童千人の身長平均が假りに一米、標準偏差が五糎とすれば、右千人中の約七百人は九十五糎乃至一〇五糎の身長を有すべく、殆ど全部が八十五糎乃至一一五糎の身長を有すと見てよいのである。この理は回歸線に於ける標準誤差についても全く同樣であつて、いま回歸線の上下に標準誤差に等しい値を取れば、その限界內に約七割が、そして標準誤差の三倍を取れば殆ど全部が含まれるのである。この場合、限界內と言ふのは回歸線を中心とした幅を指すのであつて、幅の廣さは標準誤差の二倍に當るわけである。前例に於ける回歸線は $56.7+0.297x$ であつたからその上下に標準誤差 2.7 を取つた幅は $56.7+0.297x\pm2.7$ である。換言すれば $y=54+0.297x$ と $y=59.4+0.297x$ なる二直線の間に挾まれた部分に當る。試みにこれを圖示すれば上の第五圖となる。例へば、物價が一六〇となれば、それに對應する最も確からしい賃銀は $56.7+0.297\times160$ 即ち一〇四・二二であるが、實際には $104.22\pm3\times2.7$ の範圍、即ち略ゝ九六乃至一一二の間なる事が推察されるのである。

第五圖



註　標準誤差は標準偏差までの大さを取り得る事に注意されたい。前例では賃銀系列の標準偏差は次の相關係數の算出に於て示される通り三・五八であつて、標準誤差は上記の如く二・七である。

第五節　曲線的回歸線

上例に於ける物價に對する賃銀の回歸線は $y=56.7+0.297x$ なる直線で示されたが、回歸線なるものが元來相關圖上の各座標點を縫ふ代表線である以上、もしこれら座標點が曲線的に並んでゐる場合には、恰も時系列に於ける曲線的傾向線と同じく、曲線による回歸線を描く必要が起つて來る。初歩の統計的解析は一般に直線的回歸線に限定されてゐるやうであるが、經濟系列には寧ろ曲線的のものゝ方が多いから、全く省略するわけには行かぬ。

曲線的回歸線によつて示される相關々係を曲線的相關々係（Curvilinear Correlation）といひ、そしてこれを求める方程式は、直線の場合と同じく、長期傾向線のそれと同樣である。例へば拋物線的回歸線は次の聯立方程式を解く事によつて求められる（一〇七頁參照）。

$$\left\{\begin{array}{l}\Sigma(Y)=na+b\Sigma(X)+c\Sigma(X^2)\\ \Sigma(XY)=a\Sigma(X)+b\Sigma(X^2)+c\Sigma(X^3)\\ \Sigma(X^2Y)=a\Sigma(X^2)+b\Sigma(X^3)+c\Sigma(X^4)\end{array}\right.$$

いまA、B、C……Jなる十名の社員の入社試驗成績（X）と、各人の毎月の販賣成績（Y）について拋物線的相關々係が成り立つものとすれば、右の方程式を利用する事によつて次の計算が行へる。

	X	Y	XY	X^2Y	X^2	X^3	X^4
A	4	5	20	80	16	64	256
B	5	4	20	100	25	125	625
C	6	5	30	180	36	216	1296
D	4	6	24	96	16	64	256
E	5	9	45	225	25	125	.625
F	6	10	60	360	36	216	1296
G	6	9	54	324	36	216	1296
H	7	12	84	588	49	343	2401
I	9	11	99	891	81	729	6561
J	8	9	72	576	64	512	4096
計	60	80	508	3420	384	2610	18708

$$\begin{cases} 80=10a+60b+384c \cdots\cdots(1) \\ 508=60a+384b+2610c \cdots\cdots(2) \\ 3420=384a+2610b+18708c \cdots\cdots(3) \end{cases}$$

(1)と(2)から

$$\begin{array}{rl} & 508=60a+384b+2610c \\ -) & 480=60a+360b+2304c \\ \hline & 28=\quad 24b+306c \cdots\cdots(4) \end{array}$$

(1)と(3)から

$$\begin{array}{rl} & 3420=360a+2610b+18708c \\ -) & 3072=360a+2304b+14745.6c \\ \hline & 348=\quad 306b+3962.4c \cdots\cdots(5) \end{array}$$

(4)と(5)から

$$\begin{array}{rl} & 1428=1224b+15606c \\ -) & 1392=1224b+15849.6c \\ \hline & 36=\quad -243.6c \end{array}$$

$$\therefore\ c=\frac{36}{-243.6}=-0.14778$$

(4)に代入すれば

$$24b=28+306\times0.14778=73.22068$$

$$\therefore\ b=\frac{73.22068}{24}=3.05087$$

このb及びcを(1)に代入すれば

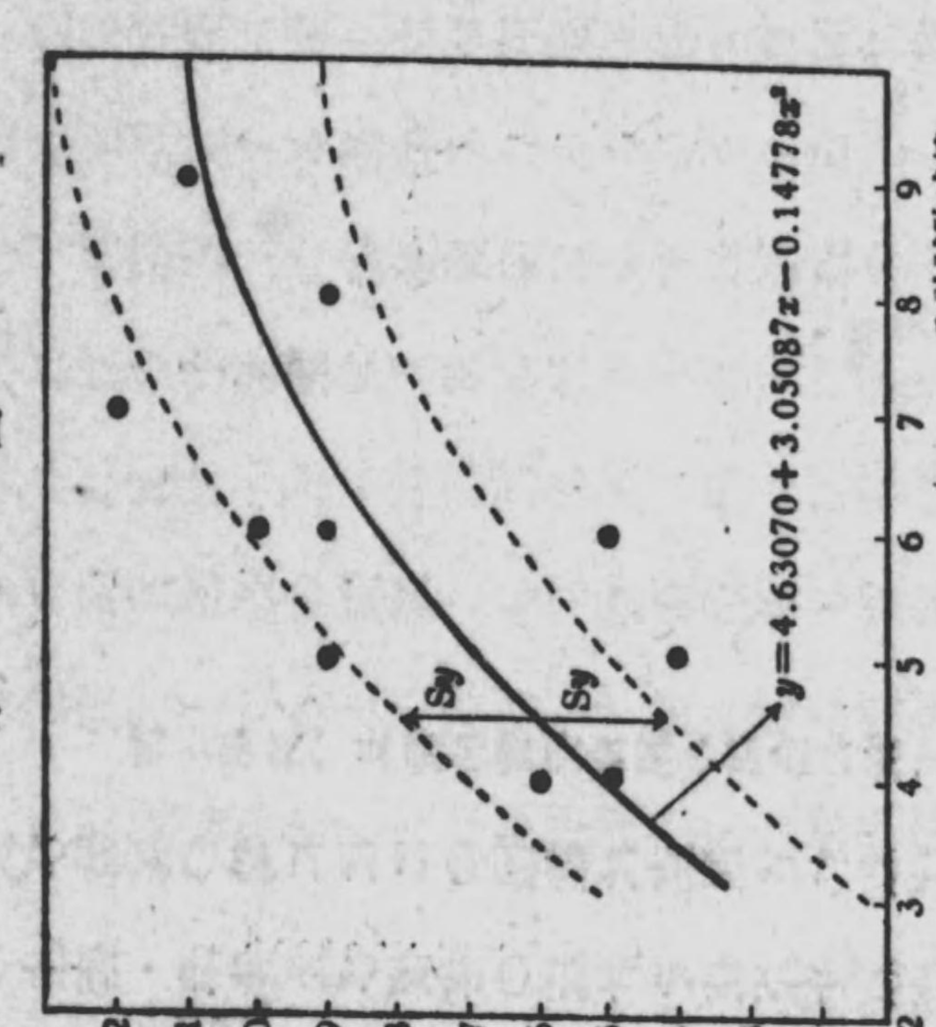
$$10a=80-60\times3.05087+384\times0.14778=46.3070$$

$$\therefore\ a=\frac{-46.3070}{10}=-4.63070$$

故に求める拋物線的回歸線の方程式は

$$y=-4.63070+3.05087X-0.14778X^2$$

第六圖



となり、第六圖の實線はこれを描いたものである。なほ曲線的回歸線についても、直線の場合と同じく、標準誤差Sを計算する事が出來る。右資料について算出したS_yは同圖に點線として描いて置いた。

理論經濟に於て必ず取扱はれる需要曲線や供給曲線などは要するに曲線的回歸線に外ならない。蓋し例へば需要曲線とは或る商品の單位價格の變化に應じて需要量は如何に變化するか、換言すれば單位價格に對應する需要量の最も確からしい大さは幾許であるかを示したものだからである。そしてその曲線の傾斜が急であるか否かによつて需要彈力性（Elaciticity of Demand）が決定されるのである。併し斯かる理論的需要曲線は單に、供給が殖えれば價格は下り、反對に供給が減れば價格は上るといふ極めて大ザツパな命題を一般的曲線によつて示したに止まり、供給がこれだけ殖えれば價格

はどれだけ下るかといふやうな具體的な答へは斯かる曲線からは求め得ない。價格變動の樣態は經濟學の中心課題であるから、吾人は各種重要商品につき具體的現實的な需要曲線を知らねばならぬ。素よりこれが爲には極めて廣範圍の生產・消費及び價格の統計を使用せねばならぬから、その作製には多大の困難が伴ふけれど、實證的經濟學の樹立及び發展の爲にはこの困難は克服されねばならぬ。この問題については「經濟統計」に於て改めて論及しよう。

註　拙稿、具體的需要曲線の導出に就て（三田學會雜誌第三〇卷第二號）參照。

第九章　相關係數

第一節　單純相關

二系列間に成り立つ相關々係を單純相關（Simple Correlation）といふ。上記の通り標準誤差の大小は相關々係の强弱を測定する基準となるものである。もし二系列間に最も强大な相關々係、卽ち數學的函數關係が存在する場合には、相關圖上の各座標點は必ず回歸線の上に位ひする事となるから、標準誤差は零となる。反之、相關々係が薄弱となれば標準誤差は次第に大きくなる。故に單に標準誤差の大小だけを見て相關々係の程度を推し得る理であるが、併しこれには大なる障碍がある。例へば前例に於ける標準誤差二・七なるものは、賃銀指數の二・七を意味し、從つて假りに通貨膨脹の結果として物價指數も賃銀指數も共に大なる數字となれば、相關々係に差が生じなくとも、標準誤差は當然大きくなるであらう。故に單に標準誤差だけで相關の程度を推す事は不可能と言つてよいのである。

この障碍を除くには標準誤差をば標準偏差に對する比率卽ち $\frac{S_y}{\sigma_y}$ として示せばよい。前例に於ける賃銀系列に就いて標準偏差（卽ち σ_y）を求めれば

	偏差 d	d^2
101	+1	1
96	−4	16
96	−4	16
95	−5	15
95	−5	15
100	0	0
103	+3	9
104	+4	16
104	+4	16
105	+5	25
100（平均）		149

から　$\sigma_y=\sqrt{\frac{\Sigma(d^2)}{N}}=\sqrt{\frac{149}{10}}=3.85$

であるから、$\frac{S_y}{\sigma_y}=\frac{2.7}{3.85}$ となる。もし全く完全な相關々係があれば標準誤差は零となるから $\frac{S_y}{\sigma_y}=0$ であり、反之、全く相關々係が無ければ S_y は σ_y に一致するから $\frac{S_y}{\sigma_y}=1$ となる。故にこれを書き代へて $1-\frac{S_y}{\sigma_y}$ とすれば、完全な相關々係は一となり、完全の無關係は零となる。換言すれば零から一までの間の數字によつて相關の程度を具體的に示し得るのである。

併し今日一般に用ひられる算式としては、

$$r=\sqrt{1-\frac{S_y^2}{\sigma_y^2}}\ \cdots\cdots\cdots\cdots(1)$$

が用ひられ、更に實際の計算に當つては、これを更に展開して得られる次の二次の何れかゞ用ひられる。

$$r=\frac{\Sigma xy}{n\sigma_x\sigma_y}\ \cdots\cdots\cdots\cdots(2)$$

(又は) $$r=\frac{\Sigma xy}{\sqrt{\Sigma x^2\Sigma y^2}}\ \cdots\cdots\cdots\cdots(3)$$

この三つの式を前例に適用して見よう。第一の $r=\sqrt{1-\frac{S_y}{\sigma_y}}$ を用ふれば $S_y=2.7$ $\sigma_y=3.87$ であるから

$$r=\sqrt{1-\frac{(2.7)^2}{(3.87)^2}}=\sqrt{1-0.489}=+0.71$$ となる。

第二及び第三を計算する爲には、次の如き計算表を作ると便利である。

年次	物價 X	賃銀 Y	x	y	xy	x^2	y^2
1	158	101	+13	+1	+13	169	1
2	146	96	−1	−4	−4	1	16
3	141	96	−4	−4	+16	16	16
4	134	95	−11	−5	+55	121	25
5	130	95	−15	−5	+75	225	25
6	137	100	−8	0	0	64	0
7	150	103	+5	+3	+15	25	9
8	158	104	+13	+4	+52	169	12
9	153	104	+8	+4	+32	64	12
10	145	105	0	+5	0	0	25
	145	100			+254	854	149

即ち第二式に代入すれば

$$r=\frac{\Sigma xy}{n\sigma_x\sigma_y}=\frac{254}{10\sqrt{\frac{854}{10}}\sqrt{\frac{149}{10}}}=\frac{254}{10\times9.3\times3.8}=+0.71$$

となり、第三式に代入すれば

$$r=\frac{\Sigma xy}{\sqrt{\Sigma x^2\Sigma y^2}}=\frac{254}{\sqrt{854\times149}}=\frac{254}{357}=+0.71$$

となり、何れを用ひるとも同じ結果となる。これは第二及び三の式は何れも第一式の變形に過ぎないからである。

第二節 多元相關と部分相關

以上に述べた相關々係は二箇の變數間に成り立つ相互の關係、即ち「單純相關」である。併し或る一つの事象が他の或る一つの事象とだけ關聯してゐるやうな場合は實際には寧ろ稀れで、殊に社會的現象に於ては甚だ多くの事象が相互に關聯し、相互に影響を及ぼし合つてゐるものである。例へば勞働者の賃銀は單に一般物價と關聯してゐるばかりでなく、勞働能率、企業利潤、勞働供給、勞働組合の勢力の大小、勞働立法その他種々の要素と關聯してゐるのであつて、これら諸要素の何れかゞ變化すれば、賃銀は必ず多かれ少かれ變化するであらう。斯く一變數 y と二つ又はそれ以上の變數 x_1 x_2 …との間に成り立つ相互的關係を多元相關（multiple Correlation）といひ、その程度を測定する係數を多元相關係數（R）といふ。併し研究に際し、x_2以下を假りに不變的なものとし、換言すればx_2以下の變數の影響は無いものとし、單にyとx_1との間の相關々係を知る必要の起る場合がある。これを部分相關（Partial Correlation）といひ、その係數を部分相關係數といふ。一般に多元相關及び部分相關は理論的にも技術的にも取扱ひ困難で、從つて初步の統計學では殆どこれに觸れないのが常である。此處では極く簡單な例によつて、その一端を示して置かう。

　農產物の收穫總量は作付段別と段當り收穫量によつて決定される。即ち收穫總量と作付段別及び段當り收穫量との三者の間に多元相關の成り立つ事は言ふ迄もなからう。次表は一八九四年から一九三〇年までの米國に於ける小麥に就ての數字である。（耕地面積の單位は百萬エーカー、收穫總量のそれは百萬ブッシェル、單位收穫量のそれは一エーカー當りブッシェル。）

年次	耕地面積	收穫總量	單位收穫量
1894	39.4	516.5	13.1
1895	40.8	569.5	13.9
1896	43.9	544.2	12.4
1897	46.0	610.3	13.3
1898	51.0	772.2	15.1
1899	52.6	658.5	12.1
1900	51.4	602.7	11.7
1901	52.6	788.6	15.0
1902	49.6	724.8	14.6
1903	51.6	663.9	12.9
1904	47.8	596.9	12.5
1905	49.4	726.8	14.7
1906	47.8	756.8	15.8
1907	45.1	763.9	14.1
1908	45.9	644.7	14.0
1909	44.3	700.4	15.8
1910	45.7	635.1	13.9
1911	49.5	621.3	12.5
1912	45.8	730.3	15.9
1913	50.2	763.4	15.2
1914	53.8	891.0	16.6
1915	60.5	1025.8	17.0
1916	52.3	636.3	12.2
1917	45.1	636.7	14.1
1918	59.2	921.4	15.6
1919	75.7	967.9	12.8
1920	61.1	833.0	13.6
1921	63.7	814.9	12.8
1922	62.3	867.6	13.9
1923	59.7	797.4	13.4
1924	52.5	864.4	16.5
1925	52.4	676.8	12.9
1926	56.4	831.4	14.8
1927	58.8	878.4	14.9
1928	58.3	914.9	15.7
1929	61.1	806.5	13.2
1930	59.1	850.9	14.4

この材料について二系列づつ取つて單純相關を求めれば次の結果を得る。收穫總量は1、耕地面積は2、エーカー當り收穫量は3の數字で示してあるから、r_{12}とは收穫總量と耕地面積との間の相關係數を意味し、r_{23}とは耕地面積とエーカー當り收穫量との間の相關係數を意味するのである。故に例へばr_{12}とr_{21}とは當然同じ事を意味する。

	收穫總量	耕地面積	單位收穫量
收穫總量		$r_{12}=0.782$	$r_{13}=0.587$
耕地面積	$r_{21}=0.782$		$r_{23}=0.016$
單位收穫量	$r_{31}=0.587$	$r_{32}=0.016$	

　さて右表によれば收穫總量と耕地面積との間の相關係數は〇・七八二であるが、この係數は決して兩者の一般的關係を反映するものとは考へられない。何となれば、原材料の示す通り、第三の要素たるエーカー當り收穫量は年によつて異つてゐるから、右の〇・七八二なる係數は單に右材料に於ける收穫總量と耕地面積との關聯を示すに止まるのである。故にもし一般にこの兩者の關聯を決定せんとする場合には、第三の要素たるエーカー當り收穫量を毎年等しいものと假定せねばならない。斯かる假定に立つて求められた係數は卽ち部分相關係數であつて、これを求める公式は次の通りである。

$$r_{12\cdot3}=\frac{r_{12}-r_{13}r_{23}}{\sqrt{(1-r^2_{13})(1-r^2_{23})}}$$

これを前例に適用すれば

$$r_{12\cdot3}=\frac{0.782-(0.587)(0.016)}{\sqrt{[1-(0.587)^2][1-(0.016)^2]}}=+0.956$$

となる。卽ち單位收穫量を不變とした場合には、收穫總量と耕地面積との間には殆ど完全な順相關のある事となるが、

これは吾人の常識から考へても當然である。それどころか吾人の常識から言へば、收穫總量を決定するものは元來耕地面積と單位收穫量以外には無いのであるから、假りに單位收穫量を一定とすれば、收穫總量と耕地面積との間には+1なる完全な順相關が無ければならない。然らば右の〇・九五六なる係數は未だ僅か乍ら誤差を持つてゐる事が判らう。この誤差は原材料に若干の誤差がある爲に生じたと見ねばならぬ。

同樣にして假りに耕地面積を一定とし、收穫總量と單位收穫量との間の部分相關係數 $r_{13\cdot2}$ を求める事も出來る。卽ち右公式に於ける1、2、3の組合せを變へればよいのであつて

$$r_{13\cdot2}=\frac{r_{13}-r_{12}r_{32}}{\sqrt{(1-r^2_{12})(1-r^2_{32})}}=\frac{0.587-(0.782)(0.016)}{\sqrt{[1-(0.782)^2][1-(0.016)^2]}}=+0.921$$

となる。これ亦常識的に言へば+1となるべきものであるから、略〻それに近い〇・九二一なる係數は大體に於て是認出來よう。最後に收穫總量を一定とし、耕地面積と單位收穫量との間の部分相關係數 $r_{23\cdot1}$ を求めれば

$$r_{23\cdot1}=\frac{r_{23}-r_{21}r_{31}}{\sqrt{(1-r^2_{21})(1-r^2_{31})}}=\frac{0.016-(0.782)(0.587)}{\sqrt{[1-(0.782)^2][1-(0.587)^2]}}=-0.879$$

となる。收穫量が一定ならば耕地面積と單位收穫量との間に反比例的關係、卽ち逆相關の成り立つ事は當然であり、從つて右の係數は可成りの誤差を含む事が察せられるが、併し兎に角吾人の先驗的結論に接近してゐる事は認められよう。

賃銀の如く、決定要素が極めて複雜な場合には、部分相關の測定は甚だ煩雜となるけれど、原理的には右と異るものではない。そして相互の關聯が極めて錯雜してゐる現象については、右の收穫現象に於けるやうに豫め先驗的結論を下し得ないから、一に斯かる計算に據る外はなく、又斯くの如き場合こそ斯かる計算が尊重されるのである。

次に多元相關係數の算出に移らう。收穫に關する前例に於ては、例へば耕地面積を常に一定と假定した條件の下に於ては收穫總量と單位收穫量との間に如何なる關係が成り立つかを問題としたが、今度は收穫總量と耕地面積とが直線的に結合されてゐるものとした場合に、これと單位收穫量との間に如何なる關係が成り立つかを問題とする。その尺度たる多元相關係數Rは次の公式から求められる――

$$1-R^2_{3(12)}=(1-r_{31}^2)(1-r^2_{32\cdot1})$$

これに前例に於て算出した結果を代入すれば

$$1-R^2_{3(12)}=[1-0.587^2][1-(-0.879)^2]=0.6554\times0.2274=0.149$$

$$\therefore\ R^2_{3(12)})=1-0.149=0.851$$

$$\therefore\ R_{3(12)}=\sqrt{0.851}=0.9225$$

同樣にして $R_{1(23)}$ も $R_{2(31)}$ も求める事が出來る。(計算すれば前者は 0.9701、後者は 0.9548 となる。諸君自ら檢算されたい。)

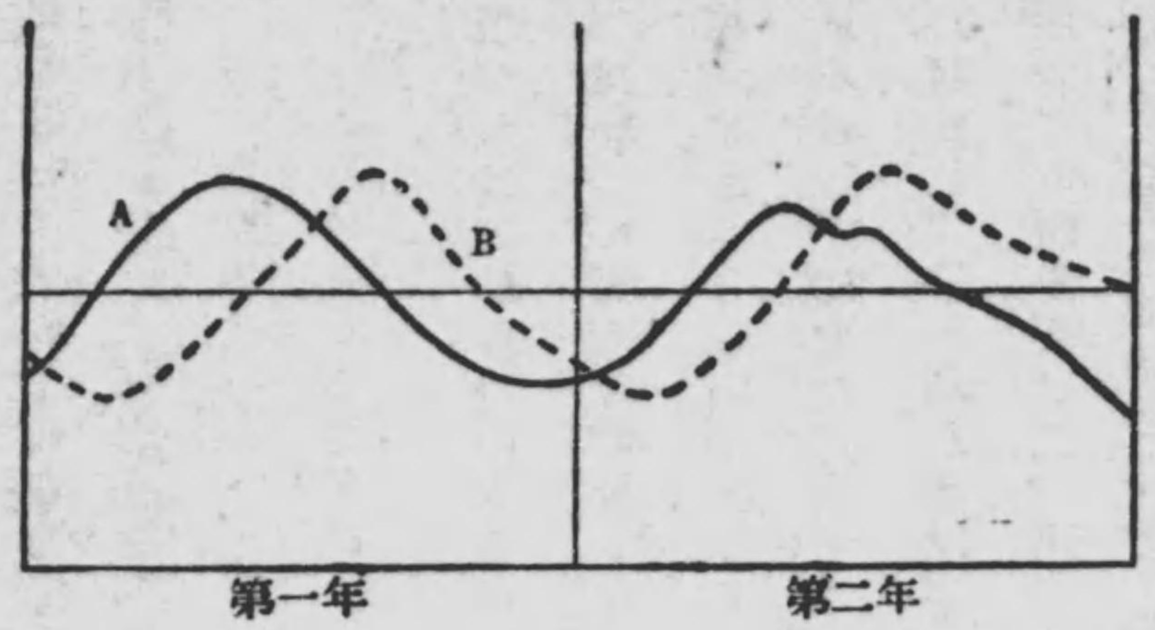
第三節　相關係數によるラグの測定

物價と賃銀、又は卸賣物價と小賣物價の如き二つの時系列間には一般に高度の相關々係が存在するもので、例へば

一般に物價の騰貴する頃には賃銀も亦これに追隨して上昇するのが原則である。併し物價が騰貴し始めれば直ちに賃銀も上昇するものではなく、實際には多少の時間的間隔を置いて上昇し始めるものであり、反對に物價が下落し始めたからといつて賃銀も同時に低下し始めるのではなく、同じく或る時間的間隔を置いて低下し始めるものである。この時間的間隔を「時の遅れ」或ひはラグ（lag）といふ。そして正確にラグを測定する事は、時系列の解析、就中景氣變動の研究に於て不可缺の要件となつてゐる。蓋しこれによつて、相關聯する二つの時系列の一方の變化は、略〻何箇月の後に他の系列に影響を及ぼし始めるかを推察し得るからである。

ラグの測定は畢竟相關係數の適用である。いま二つの時系列を$X_1 X_2 \ldots X_n$及び$Y_1 Y_2 \ldots Y_n$とすれば、既に説明したやうな單なる相關係數を算出するには、$\begin{array}{|ll|}\hline X_1 & Y_1 \\ X_2 & Y_2 \\ X_3 & Y_3 \\ \vdots & \vdots \\ X_n & Y_n \\ \hline\end{array}$と並べて計算した。Xを物價、Yを賃銀、1、2、3…nを一月、二月、三月…n月とし、假りに賃銀は二箇月のラグを以て物價に追隨するものとすれば、X_1の影響は、Y_3に於て、X_2の影響はY_4に於て、X_nの影響はY_{n+2}に於て現はれる筈である。して見れば相關係數の算出に際し、右の順序を變へ、X系列を二箇月ずらして$\begin{array}{|ll|}\hline X_1 & Y_3 \\ X_2 & Y_4 \\ X_3 & Y_5 \\ \vdots & \vdots \\ X_n & Y_{n+2} \\ \hline\end{array}$と並べれば、兩系列の間により大なる相關係數が求められるであらう。勿論賃銀が何箇月のラグを以て物價に追隨するかは豫め判つてゐないから、吾人は右の順序を逆に行つて、相關係數からラグを決定する外はない。即ちXとYとの幾多の組合せを作つて、その都度相關係數を算出し、最も大なる相關係數の得られる組合せを取つてそのラグとするのである。右の例に於ては先づ一箇月のラグを假定して、$\begin{array}{|ll|}\hline X_1 & Y_2 \\ X_2 & Y_3 \\ X_3 & Y_4 \\ \vdots & \vdots \\ X_n & Y_{n+1} \\ \hline\end{array}$と組合せて$r$を求め、次には二箇月のラグを假定して$\begin{array}{|ll|}\hline X_1 & Y_3 \\ X_2 & Y_4 \\ X_3 & Y_5 \\ \vdots & \vdots \\ X_n & Y_{n+2} \\ \hline\end{array}$と組合せて$r$を算出し、更に三箇月、四箇月等と假定してそれらの組合せの下に各〻のrを求め、是等幾多のrのうち、例へば二箇月のラグを假定した場合のrが最大の値を示すならば、賃銀は二箇月のラグを以て物價に追隨すると結論して略〻誤りが無いのである。

第七圖



斯くラグの決定には兩系列の幾多の組合せを行つてその都度rを算出せねばならぬから、若し豫めラグの大體の見當がつかなければ幾回となく計算を行はねばならず、到底その煩に堪へないであらう。そこで豫め例へばラグは略〻八箇月位であらうと、見當をつける事が出來れば、その前後の數箇月、例へば七箇月、八箇月及び九箇月の三つ位の組合せを行つて、この三つのrを比較して正しいラグを決定する事が出來よう。かやうな見當をつけるには、兩系列を同一グラフに描いて見るのがよい。上圖の如きグラフが得られゝば、B系列は略〻四箇月乃至六箇月位のラグを以てA系列に追隨する事が推察出來るから、その前後數箇月の組合せを作ればよからう。ラグの決定に於て、もし必ずしも精密を必要としないならば、相關係數による代りに、單にグラフを一覧して足りる場合が少くない。尤もこれが爲には兩系列が共に比較的に規則正しい曲線を示す場合に限られる。蓋し或る箇所(時點)では三箇月、或る箇所では八箇月といふやうに、時點によつて著しくラグが違へば、グラフを一覧した丈けでは到底見當のつかないものである。

經濟統計學の主要な一課題として統計的景氣豫測法なるものがある。その中心をなすハーヴァード研究所の方法は、投機・商況・金融の三曲線の相互のラグから、

景氣の近い將來の變化を豫測せんとするもので、相關係數の最も實用的應用の一例と見られる。（景氣豫測法については別の「經濟統計」に於て論及する。）

第四節　相關係數の妥當性

二つ又はそれ以上の異る統計系列の間に成り立つ相關々係の程度は相關係數によつて測定されるが、その使用に際しては多分の警戒を必要とする。それは假令高い相關係數が得られても、それが直ちに事實系列間に相互的關聯ある證據とはならないといふこと、況や因果關係ある證據とはならないといふことである。曾て著名な一統計學者は濠洲の或る地方では鸛の飛來する季節になると住民の間に出產率の高まる傾きある事を、兩者間の相關係數から結論した。これは決して彼が鸛が赤坊をくはへて來るといふ西洋の傳說を證明する爲ではなく、事實は反對に、もし相關係數に不用意に信頼を置いて、高い相關係數は現象間の必然且つ直接の關聯を反映するものと解釋するならば、かやうな莫迦らしい傳說をも肯定せねばならぬといふ事を警告する爲だつたのである。右の例は恐らく次の如く解釋するのが正しからう。即ち一國の出產率には季節的變動があり、(註) 濠洲の該地方では偶々出產率が高まる季節が鸛の飛來する季節と一致するに過ぎない、と。

註　我國では出產率は一月乃至三月の頃が一般に高い。これは結婚が三月前後に最も多い結果と思はれる。即ち一般に受胎は結婚直後に多いから、十箇月の懷姙期間を考慮すれば、一月乃至三月頃に出產の多いのは當然であらう。尤もこの說以外に、人間に於ても春季は性慾增進期に當り、從つてその頃受胎が多い、との說もある。

右例は偶々二現象の季節變動が一致した結果生じたものであるが、同樣の事が例へば循環運動或ひは長期傾向の一致した場合についても頻々として發生するのである。工場災害數と金利との間には屢々高い相關係數が算出されるが、これは兩者の循環運動の一致する結果である。即ち景氣上昇期には生產額を增加する爲に工場は勞働時間の延長や新規職工の傭入れなぞを行ふから各種の災害の增加するのは當然であり、又同じ景氣上昇期には一般に資本に對する需要が增加するから金利の高まるのも亦當然である。この場合、災害數と金利とは別に直接の關係に在るわけではなく、單にこの兩者が景氣變動なる廣大な運動によつて規制されてゐる丈けの事である。更にまた長期間に就て見れば、出生率と金利との間には可成り高い順の相關係數が現はれる。これは兩者が共に時代と共に低下する傾きあるが爲であるが、併し決してその間に直接の關聯があるものではない。出生率の低下は廣義の文化發展の必然的產物であり、金利の低下は資本蓄積の增大の不可避的結果たるに過ぎない。

註　出生率低下の問題については、「現代の經濟」第二卷第一號「出生率低下の原因と影響」を參照されたい。

斯く相關係數が必ずしも現象間の眞の相互關係を示すものでないとすれば、更に進んで、單に高い相關係數が求められたからと言つて、直ちに一方が原因で他方が結果であると推斷してならぬ事は明かであらう。成るほど若し現象間に因果關係があれば、原因たる現象の變化は必ず、結果たる現象の變化を伴ふから、その場合の相關係數は必ず極めて高いに決つてゐる。然るにその逆は必ずしも眞理ではない。高い相關係數も前例の如き事情の下に於ては何等の直接且つ合理的な關聯を反映するものではなく、單に間接の又は偶然の一致を反映するに過ぎない。素より高い相關係數は大多數の場合、直接又は間接の關聯ある場合に限られる。併しその關聯が直接か又は間接かは、單なる相關係

數の上からは何等知る事は出來ないのである。換言すれば高い相關係數が得られたならば、現象間に因果關係が存在するのではあるまいかとの一應の想像を廻らす事は差支へなく、否必要でもあるが、それに終局的結論を下すのは最早や統計的事實ではなくして、實に吾人の合理的思惟の力である。科學の窮極の目的たる因果關係の決定に於て、統計的方法は極めて有力な手段たる事は否定し得ないが、併し最後の最も重要な段階に於ては統計的方法は既にその偉力を喪失してゐることを忘れてはならない。そして合理的思惟とはより具體的に言へば理論に外ならぬ。即ち右の最終段階こそ、事實が終つて理論が始まる限界とも言へよう。理論のみを尊重して事實を無視するのは空中に樓閣を築くの類ひであり、反之、事實のみに偏して理論を忘れるのは佛作つて魂を入れぬに等しい。私は本講座の第一章第三節に經濟學と統計學との關聯を述べた際、統計的事實は理論の前提であり、經濟學の基本的方法は依然として演繹法たる可き旨に言及したが、その眞の意味はいま上に述べたところによつて明かになつたと思ふ。

第十章　統計學の確率論的基礎

私は緒論に於て、方法論學派の理論的基礎が確率理論にある事を一言した（一七頁）。初等統計學に於て深くこの問題を説明する事は不可能であるが、出來るだけ平易にその概略を傳へよう。本文に入るに先だち、抑も確率とは何かを明かにせねばならぬ。

「太陽は東から昇る」といふ事は吾人の知識から合理的に歸結される事柄で、これを「必然事象」といふ。反之、銀貨を投げて表が出るか裏が出るかは吾人の知識では確實には言へない事で、これを「偶然事象」といひ、偶然事象の出現する確からしさの數字的表現を確率（Probability）といふ。社會事象に就ては吾人は生起の眞の因果關係を知り得ないから、社會事象は總べて必然事象ではなくて偶然事象である。確率の概念が社會事象の研究に不可缺なるはこの理由に基く。

擲錢の結果は表か裏かゞ出るに決つて居り、且つ兩者の可能性は等しいから、表の出現する確率は$\frac{1}{2}$であるといふ（裏も亦然り）。賽を投げて例へば2の目の出る確率は$\frac{1}{6}$である。蓋し最も公平な場合には、或る目は六回に一回現はれる筈だからである。一般的に言へば、或る事象出現の可能な場合がa、不可能な場合がbだとすれば該事象出現の確率は $\frac{a}{a+b}$ である。投賽に於ては $a=1,\ b=5$ となる。

第一節　二項分布

一枚の銀貨を投擲したときは表が出るか裏が出るかに決つてゐるが、その可能性は全く等しいから、出現の確率は共に$\frac{1}{2}$である。次に二枚の銀貨を同時に投擲したとき出る表と裏の組合せは

表	表	裏	裏
表	裏	表	裏

の四通り、即ち四回のうち一回は二枚とも表、二回は一枚表一枚裏、一回は二枚とも裏であるから、出現の確率は$\frac{1}{4}$、$\frac{2}{4}$及び$\frac{1}{4}$である。同様に三枚の銀貨を同時に投擲したときは

表	表	表	裏	表	裏	裏	裏
表	表	裏	表	裏	表	裏	裏
表	裏	表	表	裏	裏	表	裏

の八通り、即ち八回のうち一回は三枚とも表、三回は二枚表一枚裏、三回は一枚表二枚裏、一回は三枚とも裏であるから、出現の確率は$\frac{1}{8}$、$\frac{3}{8}$、$\frac{3}{8}$、$\frac{1}{8}$である。かやうに次第に銀貨の數を增せば表裏の組合せは次第に複雜となるのであつて、これを表示すれば次の如くなる。

枚數	組合せの數		計
1	2	1 : 1	2
2	3	1 : 2 : 1	4
3	4	1 : 3 : 3 : 1	8
4	5	1 : 4 : 6 : 4 : 1	16
5	6	1 : 5 : 10 : 10 : 5 : 1	32
6	7	1 : 6 : 15 : 20 : 15 : 6 : 1	64
7	8	1 : 7 : 21 : 35 : 35 : 21 : 7 : 1	128
8	9	1 : 8 : 28 : 56 : 70 : 56 : 28 : 8 : 1	256

例へば八枚の場合には、表裏の組合せは九通り、即ち全部表（一回）、七枚表一枚裏（八回）、六枚表二枚裏（二八回）、五枚表三枚裏（五六回）、四枚表四枚裏（七〇回）、三枚表五枚裏（五六回）、二枚表六枚裏（二八回）、一枚表七枚裏（八回）、全部裏（一回）といふ事である。從つてその各出現確率は$\frac{1}{256}$、$\frac{8}{256}$、$\frac{28}{256}$、$\frac{56}{256}$、$\frac{70}{256}$、$\frac{56}{256}$、$\frac{28}{256}$、$\frac{8}{256}$、$\frac{1}{256}$である。

この確率は實は$\left(\frac{1}{2}+\frac{1}{2}\right)^n$なる二項式を展開すれば求められるのである（$n$は枚數）即ち

$$\left(\frac{1}{2}+\frac{1}{2}\right)^1=\frac{1}{2}+\frac{1}{2}\cdots\cdots(\text{分子},\ 1:1)$$

$$\left(\frac{1}{2}+\frac{1}{2}\right)^2=\left(\frac{1}{2}\right)^2+2+\frac{1}{2}\times\frac{1}{2}+\left(\frac{1}{2}\right)^2\cdots\cdots(\text{分子},\ 1:2:1)$$

$$\left(\frac{1}{2}+\frac{1}{2}\right)^3=\left(\frac{1}{2}\right)^3+3\times\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)+3\times\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^3\cdots\cdots(\text{分子},1:3:3:1)$$

となり、各分子が出現可能回數を示す。一般式で表せば

$$\left(\frac{1}{2}+\frac{1}{2}\right)^n=\left(\frac{1}{2}\right)^n+n\times\left(\frac{1}{2}\right)^{n-1}\left(\frac{1}{2}\right)+\frac{n(n-1)}{1\cdot2}\left(\frac{1}{2}\right)^n\left(\frac{1}{2}\right)^2+\cdots\cdots$$
$$+\frac{n(n-1)(n-2)\cdots(n-m+1)}{1\cdot2\cdot3\cdots\cdots m}\left(\frac{1}{2}\right)^{n-m}\left(\frac{1}{2}\right)^m+\cdots\cdots$$

と展開すれば、n枚投擲の場合の出現可能回數が判るのである。

上例に於ては出現確率が$\frac{1}{2}$の場合を取扱つたが（擲錢ならば表又は裏の出現確率は共に$\frac{1}{2}$である）、例へば賽

を振つて2の目の出る確率は$1/6$、出ない確率は$5/6$である。これを求めるには$\left(\frac{1}{6}+\frac{5}{6}\right)^n$を展開すればよく、一般に出現の確率が$p$、非出現の確率が$q$なる場合には$(p+q)^n$を展開すればよい。即ち

$$(p+q)^n=p^n+np^{n-1}q+\frac{n(n-1)}{1\cdot 2}\cdot p^{n-2}q^2+\cdots+\frac{n(n-1)(n-2)\cdots(n-m+1)}{1\cdot 2\cdot 3\cdots\cdots m}\cdot p^{n-m}q^m+\cdots\cdots$$

斯く事象出現の確率が豫め知られてゐる場合には上記の如き二項式を展開する事によつて最も可能的な出現回數を求める事が出來る。斯かる出現度數の分布狀態を二項分布といふ。そして擲錢の場合の如く、pとqとが相等しいときは、前例の示す通り、必ず或る一點を中心として左右に對稱的に分布してゐるから、これを對稱分布（Symetrical Distribution）又は正常分布（Normal Distribution）といひ、それを描いた曲線を正常度數曲線（Normal Frequency Curve）といふ。然るに投賽の場合の如く、pとqとが相等しくないときは、右又は左に偏つた分布狀態を示すのであつて、これを非對稱分布といふのである。併しこの場合にもp又はqが餘り小さな數でない限り、もしnを非常に大きくすれば、正常分布に著しく接近する。諸君自ら例へば$1000\left(\frac{3}{10}+\frac{7}{10}\right)^{20}$を展開して證明されたい。

第二節　正常曲線

實際に得られる度數曲線は或る度數を中心とし左右に多かれ少かれ偏倚を示す非對稱曲線であるが、偏倚の甚だしくないものに就ては正常曲線を當嵌める事が出來る。正常曲線は凡ゆる度數曲線の典型的なもので、それに關する幾多の理論が展開されてゐるから、もし斯かる曲線が當嵌められるならば、理論的取扱ひが極めて容易となるのである。

尤も正常曲線の理論を充分に理解する爲には至難の數學を用ひねばならぬから、こゝでは纔かにその輪郭を述べるに止める。

正常曲線は誤差曲線、確率曲線、正常確率曲線或ひはガウス曲線、時にはラプラス曲線なぞと呼ばれ

$$y=\frac{N}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}$$

なる方程式を以て表はされる。この方程式のyは縱座標、xは橫座標、πは略ゝ三・一四一六なる圓周率、eは略ゝ二・七一八なる自然對數の底、Nは度數分布に於ける總度數、σはその標準偏差である。πとσとは如何なる場合にも不變の常數（Constant）であり、Nとσとは曲線が異れば自ら異る常數、即ち媒介常數（Parameter）である（直線方程式$y=a+bx$又は拋物線方程式$y=a+bx+cx^2$のa、b、c等は媒介常數である）。

さて略ゝ正常分布を示す度數分布曲線に正常曲線を當嵌めるには二つの方法がある。一は面積法であり、他は縱座標法である。同一資料にこの二つの方法を施せば多少異つた結果を示すが、大體に於ては相等しいから、こゝでは前者、即ち面積法のみを述べよう。

正常分布は對稱分布であるから、算術平均・並數及び中位數は何れも相等しく、何れも曲線の頂點の示すx座標に一致する。この平均を中心として左右に標準偏差σに等しい距離を取れば、その範圍即ち$Ma\pm\sigma$と曲線とによつて包まれる面積は全度數の六八・二六％に該當し、標準偏差の二倍を取れば九五・四六％、三倍を取れば九九・七四％に該當する（九一頁參照）。故にもし左又は右の一方にσだけを取ればその面積は六八・二六％の半分、即ち三四・一三％、