統計数理に関する所感

佐々木達治郎

Considerations on the Theory of Statistcs

André G. Laurent の言を借りれば，数理統計学では数はそれが何を数えたものであるかに全く無関係に取扱われ，資料の物質的性質に関係なく作られた全く形式的な規約である．然し実際問題に於ては資料の性質を考え且つその資料の得られた世界の状態を充分に考慮してこの規約を適用しないと誤つた結果に到達するであらう．統計理論はこのままの姿でよいであらうか？　これに就て少しく論じて見たい．

数はベクトルで表わされる．然し力学に於て取扱われる量はこの外にスカラー及びテンソルがある．ベクトルの考えは一番考え易いので今迄の統計理論はベクトルの統計理論である．スカラー及びテンソルに就ての統計に於てはベクトル統計理論より変数変換によつて得られた例が 2，3 あるが統一された理論は見当らない．今これ等の例に就て述べて見よう．

直線はベクトルであり，面積は直角に交る二つのベクトルの積であるから廻転ベクトルであつて面に直角である．又体積は面積とこれに直交する稜との積であつて，これら二のベクトルは平行であるから体積はスカラーである．スカラーの統計理論はないが，粒子の大さの統計分布は Gibrat の法則に従うことが知られて居るからスカラー統計に於ては変数を $a\log x+b$ の形で表わし得ることが想像される．茲に x は粒子の大さである．

次に確率 p は数の比で表わされるからスカラーである．故に確率 p の統計に於ては変数は $\log p$ の形になると想像される．今電気通信に於けるエントロピー $H=-\sum p_i\log p_i$ に就て考えて見ると $\sum p_i=1$ なる条件があるから $H=-\overline{\log p_i}$ 即ちエントロピー H は $-\log p_i$ の平均値として表わされることが判る．この考が電気通信に於て重要なのである．

これ等の例はスカラー統計に於ては，その変数の対数の一次式によつてベクトル統計に変換し得ることを示して居るが，このことは一般に成立するであらうか？　考究を要するのである．

所得額と所得人員との間には Gibrat の法則が成立するという説がある．然るときは所得額もスカラーと考えてよいのであらうか？　若し，これが真であれば所得額の対数の平均値を M とするとき，$M=1/N\sum_{i=1}^{m}n_i\log x_i$，茲に N は全人員，となり $\left(\prod_{i=1}^{m}x_i^{n_i}\right)^{\frac{1}{m}}$ 即ち x の幾何平均が望ましい平均値ということになる．

このように変数の対数をとる例は他に幾らもある．これを皆スカラーと見てよいのであらうか？例えば

Domar の潜在的社会生産力 σ は次の如く定義せられて居る．$\sigma=\dfrac{dP}{dt}\Big/I$，茲に P は生産能力，I は1年間の投資額である．限界貯蓄性向を μ とすれば所得 Y を P に等しいものとして計算し $dI/I=\sigma\mu\,dt$　$\therefore\ \log I=\sigma\mu t+\text{const.}$ となる．若し I に就て統計をとるとすれば $\log I$ の形の変数を用いた方が有意義の如く考えられる．

又物理学の熱学に於てエントロピーなる概念がある．これを統計力学に於て論ずれば準静的過程に於てエントロピー S は

$$S=k\log(\Omega/N!)$$

で表わされる．茲に Ω は状態密度であつて，これは次の量の函数である．一つの系のエネルギー E，体積 V，粒子数 N，外部変数 x，k はボルツマン常数．ここにも対数の形が表われて居る．

さきに挙げた電気通信のエントロピーは，統計力学のエントロピーと同形であつて，或偶然量 x の確率分布密度 $f(x)$ はこの量の知識を代表する．その知識の精度を測る尺度は

$$H=\int_{-\infty}^{\infty} f(x)\log f(x)\,dx$$

で表わされこれを Informmation という．統計理論に於てこの H なる量に就て論議されて居るのは衆知の通りである．

上述の如く我々の前には変数又はその函数の対数が多く現われて居るが，スカラーと関係があるのであらうか？ 兎に角スカラー統計の一般理論を考究すべきであると考える．

次にテンソル統計の例に就て述べる．相関係数 r は一般にテンソルである．Fisher は $z=1/2\log 1+r/1-r$ なる変換によつてベクトル統計に変換して居る．z なる函数は複素変数 $\xi+i\eta$ なる平面に於て ξ 軸の -1 と $+1$ との間にあるスリットの上のポテンシャルであつてポテンシャル函数は η 方向に 2π の週期を有する．このスリット上に於て $+r$ と $-r$ とに二点があり $|r|\leqq 1$ である．故に $+r$ があれば $-r$ があるテンソルの性質を有する変数 z を用ひて変換したものである．

これはテンソル統計の一例であるが，一般理論は未だないようである．これも考究を要する問題であると考える．

次に多変数の場合には各変数が同一な性質を有する数である場合は少い．かかる異質な数の間の関係では物理学或は工学ではこれ等の変数に各々変換を行つて同一性質の数として，それ等の間の関係に就て議論するのが通則であつて，さもないと誤つた結果に到達する．統計に於てこのようなことが行われて居るであらうか？

社会学の資料から得られた数では資料の物質的性質により，上述の如き明確なる性質を有する数が考えられるであらうか？ 若し判つて居るとすれば各々変換を行うべきである．さもないと互に加算を行うことが出来ず，まして掛算は出来ない．故に数の性質を考えないでこれ等の数の算術を行つたとすれば当然間違つた結果に到達する．

多変数の場合には各々の変数に変換を行えばその頻度も変る．それ故調査より得られた頻度にも適当な修正を行うべきである．これ数量化理論の性質に類似であると考える．この為め変数変換を実証的に行ひ妥当な修正を行うべきであるが，非常な計算力を必要とするのであつて，計算法を研究し計算機械を改良して統計理論の進歩発達を計るべきであると考える次第である．

『十 周 年 に 当 り て』

松 下 嘉 米 男

To the Tenth Anniversary of the Institute

昭和 19 年，当研究所が創立されてより今年はその十周年を迎えるのであるが，今その過ぎ来りし跡を顧みれば，我々研究所と共に歩んで来たものにとつて誠に感慨新たなるものがある．次に今迄の概観を述べ，併せて今後の発展方向にも言及したいと思う．

先づ創立当初の半年ばかりは，所長以下所員三名，事務員若干名であつたが，漸次人員を増加して翌 20 年4月頃迄には所員7名，助手7名程になつた．又建物は戦時中ではあり新設されることもなく上野の学士院の一部を借りていたのであつたが，空襲の烈しくなるにつれて，翌 20 年3月に，所長と所員2名，助手2名を残しあとのものは信州飯田に疎開した．東京残留組は同時に，小