アー・イ・イェジョフ著『分布系列の平滑化と計算[1]』

А. И. Ежов, Выравнивание и Вычисление Рядов Распределений, (Госстатиздат, 1961), 336 стр.

竹内　清

1

本書は，分布の若干の特性値，たとえば，変量の算術平均とか総度数，または変動の上限および下限等の情報を援用して，全体の分布を還元する問題に主眼がおかれている。このことは経済統計の分野ではしばしばぶつかる問題である。たとえば，もとの個票とか詳細な統計表がなく，対象集団の算術平均とか，変動の上限と下限しか与えられていない場合には，本書で考察されている方法は，実際的価値も高く，分布系列の加工ならびに分析に従事する実務家のためにもその利用価値は大きいであろう。

さて本書の構成は，第Ⅰ部と第Ⅱ部からなっているが，第Ⅰ部は2つの章からなり，第Ⅱ部の導入部となっている。第Ⅱ部は10個の章からなっているが，本書全体はつぎのような構成になっている。

(1) 筆者の入手した本書は，森田優三教授が1962年4月，国連の社会経済理事会の国連統計委員会に日本政府代表として出席された折，ソ連政府代表として出席の Ежов 教授から森田教授を通じて筆者に贈られたものである。本書の書評の契機を与えられた Ежов 教授ならびに森田教授に心から感謝の意を捧げる次第である。

2

第1章変動系列では，変動系列の概念，それの処理法と関連しての分布の特性値その他の基礎が述べられる。そこでは——本書全般を通してであるが——社会経済の領域からの具体例が豊富に例示されているので，読者にとって理解しやすく，また興味をよぶであろう。

第2章は，分布の平滑化であるが，経験的な度数系列に当てはめるべき理論的曲線としては，2次の放物線がほとんどすべての例において採用される。統計学においてより一般的な分布である，正規分布とかポアソン分布については，ほんの僅かばかりふれてあるだけである。分布の平滑化の方法としては，通常最も広く利用されている古典的な最小二乗法と，実際にあまり利用されていないモーメント法が第2章の終りの部分で検討される。

ところで本書で考察している分布の平滑化と計算は，1変量の場合を対象としており，基本的には多変量を対象とする回帰分析を行なってはいな

い。[(2)] したがって度数を変量の函数として，その平滑化と計算を考察しているわけである。変量を X，その度数を Y とした場合，$Y=f(X)$ として問題を考えているわけである。変量 X が離散的である場合と連続的である場合に分けて問題を考えているが，考え方の基本そのものは同じである。実際の計算に当たっては，一定の変数変換を行なった上で，極めて簡単な計算公式を与えている。

既述のごとく，本書の第Ⅱ部は，著者によって研究された基本的な問題に当てられている。すなわち，研究対象となっている統計集団の経験的分布についての詳細なデータが欠如しており，その全体について，たとえば，算術平均とか総計とか，変動の上限と下限等のような一般的な情報しか利用できない場合に，その特性値の可能な分布をどのようにして得るか，という問題がそこで主として考察される。

第３章の初めで，分布の平滑化と計算についてふれ，上述のように算術平均とか総計とか，変動の上限と下限等の一般的な情報しか利用できない場合には，すなわち，系列の実際の度数がない場合には，計算された変数の値とそれとを比較できないので，最小二乗法は適用できない。したがって，そのような場合，計算される分布の対応するパラメーターを見出すためには，別の基礎によらなければならないことになる。

そこで著者は，そのような基礎として，全体についての一般的情報——その総度数と算術平均，それにまた変量の最初の値と最後の値，それと分布の形——をもつだけでよいであろうとする。本章では，この条件のうちの最後の二つ，すなわち，分布の形と変量の最初と最後の値を検討する。

---

(2) 最小二乗法の取扱いも最も単純なものである。経済分析で最も基本的に利用される回帰分析における最小二乗法の一般については，次書は極めてすぐれている。

Ю. В. Линник, Метод Наименьших Квадратов и Основы Теории Обработки Наблюдений изд. второе, дополненное и исправленное, 1962.

ここでは，説明度数と被説明変数の両方に誤差のある場合の，A. Wald の直交回帰の方法にもふれている。

分布の形の選択の問題について，著者はつぎのように考える。与えられた分布の形を正しく選択するための基礎としては，まず現象の本質，全体の構造の形成の合法則性，度数と変量の間の依存関係についての知識であるとして，1次函数，2次放物線，双曲線について簡単にその性質を述べる。

また変量の最初の値と最後の値を正しく選択するための重要な規準として，著者は，実際的な試行錯誤と研究対象となっている現象の本質に関する知識をあげている。（これと関連して第Ⅸ章の初めでこの問題についての補足が行なわれている)。

この章の後の部分で，特別な分布として，ポアソン分布と二項分布が具体例を用いて述べられている。ただし，ポアソン分布の前提条件とか，どのような場合に応用できるかにはふれていない。またポアソン分布のつぎに2項分布が述べられているが，この両者の理論的な関係についてはふれていない。ある分布を適用する場合の前提条件を十分に注意しておかないと，特に初心の実際家がそれを適用する際は形式的になり，実際の場合かなりの危険も予想されることはいなめないであろう。

第Ⅳ章では，1次函数による分布の計算が述べられるが，その考え方は，以下の展開の基礎になるものである。ここで著者のオリジナルな考え方の緒口を見出すことができるであろう。

ここでは，本書の中心的な役割を果たす第Ⅴ章をやや詳細に紹介することにしよう。

まず総度数 Σy，総計，算術平均 $\overline{\mathrm{X}}$，そのほか変量の下限 a と上限 b に関する情報だけが利用できる場合について，2次放物線での分布の計算を考える（基本的な離散的変量の場合について)。

2次放物線 $y={}_2a_0+{}_2a_1x+{}_2a_2x^2$ の場合，最小二乗法における正規方程式はつぎのように与えられる。

$$ {}_2a_0S+{}_2a_1\sum_{-m}^{+m}x_i+{}_2a_2\sum_{-m}^{+m}x_i^2=\Sigma y $$

$$ {}_2a_0 \sum_{-m}^{+m} x_i + {}_2a_1 \sum_{-m}^{+m} x_i^2 + {}_2a_2 \sum_{-m}^{+m} x_i^3 = \Sigma xy $$

$$ {}_2a_0 \sum_{-m}^{+m} x_i^2 + {}_2a_1 \sum_{-m}^{+m} x_i^3 + {}_2a_2 \sum_{-m}^{+m} x_i^4 = \Sigma x^2 y \quad . $$

上式で

$$ x_i = X_i - \frac{a+b}{2} $$

$$ S = b - a + 1 = 2m + 1 \quad . $$

x は

$$ -m,\ -(m-1), \cdots,\ -1,\ 0,\ +1, \cdots +(m-1),\ +m $$

という値をとることになる。

ところで

1) x の奇数ベキ項の和は 0 である。

$$ \sum_{-m}^{+m} x = \sum_{-m}^{+m} x^3 = 0 \quad . $$

2) $-m$ から 0 までと 0 から m までの合計は，$-m$ から -1 までと $+1$ から $+m$ までの合計に等しい。

3) 偶数ベキ項の和によって，すなわち

$$ \sum_{-m}^{+m} x_i^{2k} = 2 \sum_{1}^{m} x_i^{2k} $$

であるから，次式が与えられる。

$$ {}_2a_0 S + 2\,{}_2a_2 \sum_{1}^{m} x_i^2 = \Sigma y $$

$$ 2\,{}_2a_1 \sum_{1}^{m} x_i^2 = \Sigma xy $$

$$ 2\,{}_2a_0 \sum_{1}^{m} x_i^2 + 2\,{}_2a_2 \sum_{1}^{m} x_i^4 = \Sigma x^2 y \quad . $$

上式を解くことによって，つぎの結果が与えられる。

$$ {}_2a_0 = \frac{\Sigma x_i^4 \Sigma y - \Sigma x_i^2 \Sigma x^2 y}{S \Sigma x_i^4 - 2(\Sigma x_i^2)^2} $$

$$ {}_2a_1 = \frac{\Sigma xy}{2\Sigma x_i^2} \quad \text{または} \quad {}_2a_1 = \frac{\bar{x}\Sigma y}{2\Sigma x_i^2} $$

$$_2a_2=\frac{S\Sigma x^2y-2\Sigma x_i{}^2\Sigma y}{2[S\Sigma x_i{}^4-2(\Sigma x_i{}^2)^2]}$$

$$i=1,\ 2,\ 3\cdots\cdots m=\frac{n}{2}\quad .$$

以上のようにして，$_2a_0$，$_2a_1$，$_2a_2$ をうるためには，つぎの情報をうる必要がある。

1)　Σx

2)　Σxy

3)　Σx^2y　.

ところで，Σy と Σxy は通常統計でえられることが多いが，Σx^2y については，通常の公表統計ではえられない。そこで Σx^2y をどのようにして推計するかが問題となる。これは一般に間接的方法によることになる。

それではどのようにしてパラメーターの推計を簡単化しているか，対称分布の場合についてみよう。

二次の放物線を次式で表わす。

$$y'={}_2a_0+{}_2a_1x+{}_2a_2x^2\quad .$$

上式を対称な分布として，$_2a_0$, $_2a_1$, $_2a_2$ を見出すには，Σx^2y または $\bar{x}^2$ の大きさを決める必要はない。問題はつぎのようになる。

$$_2a_0+{}_2a_1x+{}_2a_2x^2$$

を一次の因数に分解することができる。すなわち，もし x_1 と x_2 が二次方程式

$$_2a_0+{}_2a_1x+{}_2a_2x^2=0$$

の根であれば

$$_2a_0+{}_2a_1x+{}_2a_2x_2={}_2a_2(x-x_1)(x-x_2)\quad .$$

または，結局つぎのようになる。

$$y'={}_2a_2(x-x_1)(x-x_2)\quad .$$

したがつて上式から $x=x_1$ または $x=x_2$ で

$$y'=0\ .$$

変量 x の根 $(-\mathrm{m}$ と $+\mathrm{m})$ は，絶対値が等しく符号が異なるので，二次方程式の根はまたつぎのようになる。

$$|x_1|=|x_2| \quad .$$

x_1 の符号を－，x_2 の符号を＋としよう。そうすると，任意の度数 y_i にたいして次式がえられるであろう。

$$y_i'={}_2a_2(x_i+x_1)(x_i-x_2)={}_2a_2(x_i^2-x_{1,2}^2) \quad .$$

度数 y_i を合計すると

$$\Sigma y'={}_2a_2\sum_{-m}^{+m}(x_i^2-x_{1,2}^2)={}_2a_2(\sum_{-m}^{+m}x_i^2-\mathrm{S}x_{1,2}^2)={}_2a_2(2\sum_{1}^{+m}x_i^2-\mathrm{S}x_{1,2}^2) \quad .$$

ただし $\Sigma y'=\Sigma y$.

したがって ${}_2a_2(2\sum_{1}^{m}x_i^2-\mathrm{S}x_{1,2}^2)=\Sigma y$.

これから ${}_2a_2$ はつぎのように容易に求める。

$${}_2a_2=\frac{\Sigma y}{2\sum_{1}^{m}x_i^2-\mathrm{S}x_{1,2}^2} \quad .$$

このようにして，対称分布にたいしてはつぎの結果がえられる。

$$y_i'=\frac{x_{1,2}^2\Sigma y}{\mathrm{S}x_{1,2}^2-2\sum_{1}^{m}x_i^2}-\frac{\Sigma y}{\mathrm{S}x_{1,2}^2-2\sum_{1}^{m}x_i^2}x_i^2$$

$$=(x_{1,2}^2-x_i^2)\frac{\Sigma y}{\mathrm{S}x_{1,2}^2-2\sum_{1}^{m}x_i^2} \quad .$$

上式において $x=x_1$ または $x=x_2$ で端の縦座標の値は 0 になる。だが，圧倒的多数の社会生活現象では，特性の与えられた具体的な値に——初めと終りに—— 0 と等しくない一定の度数が対応する。そこで，両端の値より絶対値が 1 だけ大きい特性値に対応する度数が 0 であると考えて，著者は対称的な 2 次方程式の根を求める。すなわち

$$x_1=-(\mathrm{m}+1)$$

$$x_2=(\mathrm{m}+1)$$

と考えるわけである。したがって

$$x_{1,2}^2 = +(\mathrm{m}+1)^2 \quad .$$

ところで $\mathrm{S}=2\mathrm{m}+1$ から $\mathrm{m}+1=(\mathrm{S}+1)/2$。したがって上式はつぎのようにも表わされる。

$$x_{1,2}^2 = = \frac{(\mathrm{S}+1)^2}{4} \quad .$$

また

$$\sum_{1}^{\mathrm{m}} x_{\mathrm{i}}^2 = \frac{\mathrm{m}(\mathrm{m}+1)(2\mathrm{m}+1)}{6} = \frac{\mathrm{S}(\mathrm{S}^2-1)}{24}$$

から，${}_2\mathrm{a}_0$，${}_2\mathrm{a}_1$，${}_2\mathrm{a}_2$ はつぎのようにも求まる。

$${}_2\mathrm{a}_0 = \frac{x_{1,2}^2 \Sigma y}{\mathrm{S}x_{1,2}^2 - 2\sum_{1}^{\mathrm{m}} x_{\mathrm{i}}^2} = \frac{3(\mathrm{S}+1)\Sigma y}{2\mathrm{S}(\mathrm{S}+2)}$$

$${}_2\mathrm{a}_1 = 0$$

$${}_2\mathrm{a}_2 = -\frac{6\Sigma y}{\mathrm{S}(\mathrm{S}+1)(\mathrm{S}+2)} \quad .$$

つぎに二次放物線について，分布の中央に関して非対称な場合のパラメーター ${}_2\mathrm{a}_0$，${}_2\mathrm{a}_1$，${}_2\mathrm{a}_2$ はつぎのようにして求められている。

$${}_2\mathrm{a}_0 + {}_2\mathrm{a}_1 x + {}_2\mathrm{a}_2 x^2 = 0$$

を解くと，根は次式で与えられる。

$$x_{1,2} = \frac{-{}_2\mathrm{a}_1 \pm \sqrt{{}_2\mathrm{a}_1{}^2 - 4{}_2\mathrm{a}_0 \cdot {}_2\mathrm{a}_2}}{2{}_2\mathrm{a}_2} \quad .$$

凸分布の頂点が分布の中央の左側にある場合，大きな根 x_2 は上式分子の平方根の前の符号は＋になる。対称的な場合と同様に，この場合にもつぎのように考える。

$$x_2 = \mathrm{m}+1 = \frac{\mathrm{S}+1}{2}$$

または

$$\frac{-{}_2\mathrm{a}_1 + \sqrt{{}_2\mathrm{a}_1{}^2 - 4{}_2\mathrm{a}_0 \cdot {}_2\mathrm{a}_2}}{2{}_2\mathrm{a}_2} = \mathrm{m}+1 = \frac{\mathrm{S}+1}{2} \quad .$$

これを，すでに２次放物線の場合の正規方程式群の最初の二つ，すなわち

$$ {}_2a_0 S + 2{}_2a_2 \sum_1^m x_i^2 = \Sigma y $$

$$ 2{}_2a_1 \sum_1^m x_i^2 = \Sigma xy $$

に含めて考える必要がある。上の最後の式から

$$ {}_2a_1 = \frac{\Sigma xy}{2\sum_1^m x_i^2} . $$

または

$$ \Sigma x^2 = \frac{m(m+1)(2m+1)}{6} = \frac{S(S^2-1)}{24} $$

から

$$ {}_2a_1 = \frac{12\Sigma xy}{S(S^2-1)} . $$

これ以下前と同様に展開して，つぎの結果がえられる。

$$ {}_2a_0 = \frac{3[(S+1)+2\bar{x}]\Sigma y}{2S(S+2)} $$

$$ {}_2a_1 = \frac{12\Sigma xy}{S(S^2-1)} $$

$$ {}_2a_2 = -\frac{6[(S-1)+6\bar{x}]\Sigma y}{S(S+2)(S^2-1)} . $$

左側に非対称な場合——社会経済現象としては稀であるが——については，つぎのようになる。

$$ {}_2a_0 = \frac{3[(S+1)-2\bar{x}]\Sigma y}{2S(S+2)} $$

$$ {}_2a_1 = \frac{12\Sigma xy}{S(S^2-1)} $$

$$ {}_2a_2 = -\frac{6[(S-1)-6\bar{x}]\Sigma y}{S(S+2)(S^2-1)} . $$

したがって，右側に非対称な場合と左側に非対称な場合の，パラメーター

の値は，$_2a_1$ は両者等しく，$_2a_0$ と $_2a_2$ は，分子の $\bar{x}$ の前の符号が異なるだけである。

ところで計算の簡単化からすると，$_2a_0$ は上式から直接求めなくとも，既述の正規方程式から，パラメーター $_2a_2$ を媒介として求めてもよい。

$$_2a_0=\frac{12\Sigma y-S(S^2-1)_2a_2}{12S}.$$

計算労力を簡単化するために，著者は，総度数を 100 にした上で，便利な計算表を作製している。すなわち，$\Sigma y=100$として，上の式をつぎのように表わす。

$$_2a_0=\frac{(S^2-1)}{12}\left[\frac{1200}{S(S^2-1)}-{}_2a_2\right]$$

$$_2a_1=\frac{1200}{S(S^2-1)}\cdot\bar{x}$$

$$_2a_2=-\frac{3600}{S(S+2)(S^2-1)}\left(\frac{S-1}{6}-|\bar{x}|\right).$$

ついで

$$K^{\mathrm{I}}_{2\partial}=\frac{S^2-1}{12},\qquad K^{\mathrm{III}}_{2\partial}=\frac{3600}{S(S+2)(S^2-1)}$$

$$K^{\mathrm{II}}_{2\partial}=\frac{1200}{S(S^2-1)},\qquad K^{\mathrm{IV}}_{2\partial}=\frac{S-1}{6}$$

とおき，$_2a_0$，$_2a_1$，$_2a_2$ をつぎのように表わす。

$$_2a_0=K^{\mathrm{I}}_{2\partial}(K^{\mathrm{II}}_{2\partial}-{}_2a_2)$$

$$_2a_1=K^{\mathrm{II}}_{2\partial}\cdot\bar{x}$$

$$_2a_2=-K^{\mathrm{III}}_{2\partial}(K^{\mathrm{IV}}_{2\partial}-|\bar{x}|).$$

著者は，S＝3,……，100 に対応するそれぞれの K の値を計算してあるので，計算は容易になる。ただし，S が偶数のときは，以上の計算結果について，$_2a_1$ を 2 で，$_2a_2$ を 4 で割ることが必要になる。

3

第６章では連続的変量の分布の２次放物線での計算が述べてある。基本的な考え方は離散的な変量の場合と同じである。ここでも対称的な場合と非対称的な場合に分けて考えている。

第７章は凸曲線と凹曲線について，離散的な変量と連続的な変量の場合を考察している。考え方の基本は前章までと同じである。

第８章では，３次の放物線での分布の計算が述べられてある。

第９章は，双曲線での分布の計算が述べられている。ここでも離散的変量と連続的変量について別個に取扱っている。

社会経済現象の系列の全体の度数は，双曲線で近似できる曲線でしばしば決定される。すなわち

$$Y = A_0 + \frac{A_1}{X} .$$

A_0, A_1 はパラメーター。この A_0, A_1 も正規方程式をもとにして，これまで見た関係を利用してつぎの結果がえられる。

$$A_0 = \frac{2\left[\Sigma\left(\frac{1}{X}\right)\overline{X} - S)\Sigma Y\right]}{S\left[(S+1)\Sigma\left(\frac{1}{X}\right) - 2S\right.}$$

$$A_1 = \frac{\left[(S+1) - 2\overline{X}\right]\Sigma Y}{(S+1)\Sigma\left(\frac{1}{X}\right) - 2S} .$$

上式では $\Sigma\left(\frac{1}{X}\right)$ があるが，著者はつぎの関係を用いて近似的にそれを導き出すことを提唱している。

$$\sum_{a'}^{b'}\left(\frac{1}{X_i}\right) \frac{(b' - a' + 2a'b')S'}{a'b'(a'+b')} .$$

ただし，a' はグループ内の最初の変量値

b' はグループ内の最後の変量値

S′ はグループ内の変量の個数。

ところで上式を使う場合，X≧5 なることが注意されている。

連続変量の場合の A_0 および A_1 は次式で与えられている。

$$A_0=\frac{2[(\log_e b-\log_e a)\overline{X}-(b-a)]\Sigma Y}{(b-a)[(a+b)(\log_e b-\log_e a)-2(b-a)]}$$

$$A_1=\frac{[(a+b)-2\overline{X}]\Sigma Y}{(a+b)(\log_e b-\log_e a)-2(b-a)}\quad .$$

なお区間 $(X_i,\ X_{i+h})$ における分布の度数 Y_i' はつぎの積分によって決定される。

$$\int_{X_i}^{X_{i+h}} YdX=\int_{X_i}^{X_{i+h}}(A_0+\frac{A_1}{X})dX=A_0(X_{i+h}-X_i)$$
$$+A_1(\log_e X_{i+h}-\log_e X_i)\quad .$$

いま

$$X_{i+h}-X_i=\Delta X$$

$$\log_e X_{i+h}-\log_e X_i=\Delta\log_e X$$

とおくと

$$Y_i'=A_0\Delta X+A_1\Delta\log_e X\quad .$$

第10章ではいくつかの基本的な統計的特性値についての計算，およびそれらの間の関係について述べてある。

非対称度については，K. Pearson 流の

$$A=\frac{\mu_3}{\sqrt{\mu_2^3}}$$

（ただし，μ_3 は平均値の周りの３次の積率，μ_2 は平均値の周りの２次の積率）を検討した後，彼の測度を提起する。すなわち，非対称度として，分布の中央（変量の初めと終りの値をそれぞれ a, b とした場合 (a+b)/2)に関して全体のうちの一方の部分の他の部分にたいする関係として，または二つの部分のうちの一方の全体にたいする関係——全体にたいする部分として——を採用する。

全体（Σy）にすいする左側の部分の度数（${}_1\Sigma y$）の割合 ${}_1\Sigma y/\Sigma y$ を α_2 で

表わすと，右側の部分の度数（$_2\Sigma y$）の全体（Σy）に対する割合は $1-\alpha_2$ となる。したがって

$$\int_{-\frac{(S-1)}{2}}^{0} f(x)dx=\alpha_2\Sigma y \qquad \text{（左側）}$$

$$\int_{0}^{+\frac{(S-1)}{2}} f(x)dx=(1-\alpha_2)\Sigma y \qquad \text{（右側）}$$

となる。2次放物線の分布の場合には，α_2 として結局つぎの結果が導かれる。

$$\alpha_2=\frac{(S-1)-3\bar{x}}{2(S-1)} \quad .$$

上式はつぎのようにも表わすことができる。

$$\alpha_2=\frac{a+5b-6\overline{X}}{4(b-a)} \quad .$$

以上から，奇数個の変量（$S=2m+1$）からなる離散的分布の非対称度はつぎのように与えられる。

$$\alpha_2=\frac{S-3\bar{x}}{2S} \quad .$$

同様に偶数個の変量をもつ分布については

$$\alpha_2=\frac{(S^2-1)-3S\bar{x}}{2(S^2-1)} \quad .$$

ついで3次の放物線で表わされる連続的変量をもつ分布の非対称度についてはつぎのように導いている。

分布の左側の部分について

$$\int_{-\frac{S-1}{2}}^{0} ydx={}_1\Sigma y=(1-\alpha_3)\Sigma y \quad .$$

右側の部分について

$$\int_{0}^{+\frac{S-1}{2}} y\,dx = {}_2\Sigma y = (1-\alpha_3)\Sigma y \quad .$$

上式から

$${}_3a_0\frac{(S-1)}{2} - {}_3a_1\frac{(S-1)^2}{8} + {}_3a_2\frac{(S-1)^3}{24} - {}_3a_3\frac{(S-1)^4}{64} = \alpha_3\Sigma y \quad .$$

$${}_3a_0\frac{(S-1)}{2} + {}_3a_1\frac{(S-1)^2}{8} + {}_3a_2\frac{(S-1)^3}{24} + {}_3a_3\frac{(S-1)^4}{64} = (1-\alpha_3)\Sigma y$$

これからつぎの結果が導かれる。

$$\alpha_3 = \alpha_2 + \frac{(S-1)^4\,{}_3a_3}{320\Sigma y} \quad .$$

上式からつぎのことが分る。凸分布では，${}_3a_3$ が正であるから α_3 は α_2 より大きいが，凹分布では ${}_3a_3$ が負であるから α_3 は α_2 より小さい。

ついで著者は，算術平均と，変量の最初の値と最後の値だけの情報があり，具体的な系列は欠如している場合における，メディアン (M_e) と算術平均 ($\overline{X}$) との間の関係，モード (M_0) と算術平均 ($\overline{X}$) との間の関係，分散 (σ^2) と算術平均 ($\overline{X}$) との間の関係，四分位数 (Q_1 と Q_3) と算術平均 ($\overline{X}$) との間の関係をそれぞれ別個に導いている。[(3)]

たとえば，2次の放物線で表わされる分布の場合について σ^2 と $\overline{X}$ の間の関係はつぎのようになる。

$$\sigma^2 = \frac{(S-1)^2\{[3S(S+3)-2]+12S|\bar{x}|\}}{60S(S+1)} - \bar{x}^2 \quad .$$

離散的変量をもった分布の場合には

$$\sigma^2 = \frac{(S-1)(S+3)+4(S+2)|\bar{x}|}{20} - \bar{x}^2 \quad .$$

第11章では，変動系列の計算方法について述べてある。

---

(3) K. Pearson は正規分布にほぼ近い分布について $M_0 \fallingdotseq \bar{x}-3(\bar{x}-M_e)$ を導いた。上式では $\bar{x}$, M_e, M_0 が一つの関係で結ばれているが，著者は、これらのそれぞれ二つの間の関係を別個に導いていることに注意。

これまでの展開でも分るように，変量の最初の値と最後の値が極めて重要な働きをしてきたが，本章の初めの部分で，それらを決定するための補足的な規準が述べられる。結局，分布の最初の値 a と最後の値 b の範囲はつぎのようにまとめられる。

左に非対称な場合

非対称度	非対称度の値 (α_2)	a の範囲	a の近似的平均値
強度	0.1−0.2	$\frac{30\bar{x}-23b}{7}<a<\frac{10\bar{x}-7b}{3}$	$4\bar{x}-3b$
適度	0.2−0.4	$\frac{10\bar{x}-7b}{3}<a<\frac{30\bar{x}-17b}{13}$	$3\bar{x}-2b$
弱度	0.4−0.5	$\frac{30\bar{x}-17b}{13}<a<2\bar{x}-b$	$2\bar{x}-b$

右に非対称な場合

非対称度	非対称度の値 (α_2)	b の範囲	b の近似値
弱度	0.5−0.6	$2\bar{x}-a<b<\frac{30\bar{x}-17a}{13}$	$2\bar{x}-a$
適度	0.6−0.8	$\frac{30\bar{x}-17a}{13}<b<\frac{10\bar{x}-7a}{3}$	$3\bar{x}-2a$
強度	0.8−0.9	$\frac{10\bar{x}-7a}{3}<b<\frac{30\bar{x}-23a}{7}$	$4\bar{x}-3a$

ついでいままで述べてきた2次放物線での分布の計算の段階を，一貫してまとめてある。さらに分布系列の計算について，具体例でいくつかの異なった方法にふれている。

本章の最後の節では，度数が滑らかに変動していない場合の取扱いとして，二つの部分に分けて計算する方法が述べられている。

最後の第12章では，正規分布の計算が簡単に述べられる。

正規分布は，算術平均 $\bar{X}$ と標準偏差 σ が決まれば決ってしまう。まず第

一はそれが一定として，ついで分布の範囲の知識によっておきかえられる。その範囲として 6σ をとれば

$$\sigma=\frac{S-1}{6}$$

または

$$\sigma^2=\frac{S^2}{36}\quad.$$

このようにして，離散的な正規分布 $\left(\sigma^2=\frac{S^2}{36}\right)$ と放物線分布 $\left(\sigma^2=\frac{S^2}{20}\right)$ の間の差がえられる。

実際には，正規分布や放物線型のように厳密に対称的な分布は，極めて稀な現象である。圧倒的多数の分布はある程度の非対称度をもって非対称型の分布をもっている。しかし非対称分布でも，ある程度，一定の近似度で正規分布によって表わされる分布も稀ではない。そこで第一次近似として正規分布で表わされる分布について，系列の変動はつぎの指標によってほぼ特長づけられる，ということを導いている。

連続的分布にたいして

$$\sigma=\frac{(S-1)}{6}\sqrt{\frac{S-1}{(S-1)-3|\bar{x}|}}\quad.$$

離散的分布にたいして

総度数が奇数個の場合

$$\sigma=\frac{S}{6}\sqrt{\frac{S}{S-3|\bar{x}|}}\quad.$$

総度数が偶数個の場合

$$\sigma=\frac{S}{6}\sqrt{\frac{S^2-1}{(S^2-1)-3S|\bar{x}|}}\quad.$$

以上のようにして本文は閉じるが，付録に自然数 $S=1,2\cdots\cdots,100$ について $S^2, S^3,\cdots, S^7$ の表，$\Sigma(S), \Sigma(S^2),\cdots\cdots, \Sigma(S^7)$ の表，正規分布の表，ポアソン分布表，e^x の表，e^{-x} の表がのっている。付録でポアソン分布表がかな

り詳しくのっている割には，本文の説明が不十分であるといえよう。

4

さて以上簡単に本書の概要をみたが，著者のオリジナルな方法——欧米の統計学の書物でわれわれのなじんでいないものであるが——は極めて高く評価されるべきであるが，なお不十分と考えられる点も若干あるので，基本的なものについてだけ以下に簡単にふれることにする。

まず第一は，本書で引用して分析を行なっている豊富な具体例については，それがどのようにして得られたのか，またそのデータの信頼度はどの程度のものか，等々いわゆる経済統計の基礎的な資料論とも関連した問題の取扱いが十分になされていない点である。したがって，データの分析法もいわゆる既述統計的な段階のものとなっている。データの性格によって，それ以降の分析法も基本的に制約される性質をもっていることを，われわれは銘記すべきである。データの性格を十分につかんでおかないと，適切な分析はできないともいえるであろう。

データの性格についての吟味が，以上のように十分行なわれていないので，推測統計的な取扱いは全く行なわれていない。そのため，本書では確率論の基本についても，また標本調査論の基本などについてもふれていない。ソ連においても，もちろん，このような問題は，社会経済統計の分野でもますます大きな意義をもってきている。文献によっては，もちろん，これらの問題にもふれている[(4)]。共産主義社会建設のために，経済研究にますます数学的方法の利用が実際的な重要性をもちつつある今日のソ連では，統計分析に当たっても，推測統計的技法の適用できる分野ではその基本は実際家にとっても必要であろうと考えられる。

つぎに本書の性格から逸脱するかもしれないが，単一変量の分布の平滑化

(4) たとえば，経済研究家のための数学の入門書としての次書にもふれてある。
А. Я. Боярский, Математика для Экономистов, 1961.

と計算だけでなく，他変量との回帰分析——統計分析ではこの方が一般に重要であるが——との関連にも言及が行なわれていると，実際家にとってはより有用であったであろうと考えられる。

前節までの概要から分るように，統計集団についての一般的な情報として，その大きさ，算術平均，変動の上限と下限，それに分布の形は，著者の展開している結果を利用するに当たっては必要となる。

ところで与えられた統計集団の特性がどのような分布の形をもつかを，上述の情報だけから予め判断するのは，なかなか難しいであろう。特に初心の実際家にとっては。もちろん，すでにふれたように，著者もある程度は若干の分布の性格について述べている。しかし実際家にとっては，ある程度の経験を積まないと，本書だけの説明では不十分であろう。分布の形の選択は，とかく主観的なものになり易い性質をもっている。この問題と，変量の最初の値と最後の値の選択の問題——これもかなり主観的なものになり易い——が，著者の方法を適用する場合の最大の制約になるであろう。これを救う道は，著者もいっているように，ある程度の試行錯誤をやって経験を積むとともに，現象の本質を深く洞察する力を高めることにあると考えられる。

曲線の当てはめの問題に関しては，変量が等間隔に測られている場合には，直交多項式の利用が有用であろう(5)。すなわち，これを利用すると，一次式で当てはめが不十分なときは，一次の係数はそのままにして，二次の係数だけを計算すればよい。もし二次式で当てはめが不十分なときは，一次と二次の係数はそのままにしておいて，三次の係数を計算すればよい。このように直交多項式を利用すると，低次の係数はそのままにしておいて高次の係数を計算することができて，計算に便利である。普通の回帰式の場合は，その都度，全部の係数を計算し直さなければならない。実際家にとっては，その結果の利用法だけを知っていれば十分であろう。

---

(5) 曲線の当てはめに関しては，たとえば，極めて簡にして要を得てすぐれた次書と比較すると，この点がよく分るであろう。
増山元三郎,『実験公式の求め方』, 1962.

推測統計的な取扱いをしていないので，当てはめた曲線が妥当かどうかの判断は恣意的なものになる危険がある。これも，たとえば，適合度の検定その他の検定法により，計算した曲線が果たして妥当なものかどうかの判断が確率論的に行なえるであろう。

以上本書を繙いた後で，考えついたいくつかの問題のうちの若干のものを簡単に述べたわけである。

もちろん，筆者の考えたことは，著者も万々承知の上であり，あえてそれらの問題を切り捨てたものであろう。すなわち，数学ならびに数理統計学の水準の高くない実務家にとっては，筆者のふれたような推測統計学的手法を取り上げると，著者の意図である数理統計学の国民経済への応用という点で，本書の実際的価値のそこなわれるのを，著者が慮ったがためであろう。

筆者も本書を読んで非常に興味を感じ，これをさらに展開してみようと考えている次第である。最後に，A. Длин の本書にたいする書評[(6)]の最後の言葉をもって，この書評の筆を擱くことにしよう。「…A. И. Ежоов の著作は，ソヴェト統計に従事する者にまた，多くの統計的研究への数学の巧妙なそして正しい利用の方法を供給し，統計の仕事をより生産的にし，一連の場合において，莫大な労力と金のかかる統計の作成から手を引かしめる。最も重要なのは——短い時間で全体の分布についての予備データをうることを可能にしていることである。」 (1963—1—12)

付　記

本稿を脱稿後，Г. Кильдишев，В. Овсиенко の本書に対する書評[(7)]を読む機会をえたので付記しておく。Вестник Статистики はソ連における統計学の代表的な月刊学術雑誌であり，Вопросы Экономики は経済学の代表的な月刊学術雑誌であるから，これらに載った書評を読めば，本書に対する評価がソ

---

(6) А. Длин, "А. И. Ежов «Быравнивание и Вычислеиие Рядов Распределений»," Вестник Статистика, 1962, No. 8, стр. 69—71.

(7) Г. Кильдишев, В. Овсиенко, "Книга о Математических Методах Обработки Статистических Данных—А. И. Ежов, Выравничие и *

連においてどのようなものであるかを十分に知ることができるであろう。Длин と Кильдишев, Овсиенко の本書に対する評価は，細部にわたっては若干の差異があるが，基本的には等しく本書を高く評価している。

* Вычисление Радов Распределений," Вопросы Экономики, No. 12, 1962, стр. 114—117.