論というものの根本性格である。その条件は，$\varphi(\mathrm{E})$ は当然Ωに属すること，次に，$\varphi(\mathrm{E})$ が未知の分布関数Ｆの良好な統計的推定（good statistical estimate）であるということである。ここに good statistical estimate ということは，一般的には次のように規定されよう。すなわち，ある小区域に $\varphi(\mathrm{E})$ が存在する確率をできるだけ大きくしたものである。

たとえば，$\mathrm{X}_1, \mathrm{X}_2, \cdots, \mathrm{X}_n$ に関して，(10) 式の成立が前提され，しかも標準偏差 $\sigma=1$ として前提されたとする。このとき，標本 $\mathrm{E}=(x_1, x_2, \cdots, x_n)$ をもととして，母集団平均値mを推定しようという問題が起こるであろう。従来のやり方は，

$$\varphi_1(\mathrm{E})=\frac{x_1+x_2+\cdots+x_n}{n} \tag{12}$$

という方法であった。しかしこれだけがただ一つの方法ではない。たとえば

$$\varphi_2(\mathrm{E})=\text{Median of }(x_1, x_2, \cdots, x_n) \tag{13}$$

とすることも可能である。そこで，$\varphi_1(\mathrm{E})$ と $\varphi_2(\mathrm{E})$ とはいずれがよいかという問題もおこる。

6.　推測統計学の諸問題

R. A. Fisher 以来，近代統計学の根本的な問題は，次の 4 つに要約される。

〔Ⅰ〕 特徴づけの問題 (Problems of specification)

〔Ⅱ〕 統計的仮説検定の問題 (Problems of testing statistical hypotheses)

〔Ⅲ〕 統計的推定の問題 (Problems of statistical estimation)

〔Ⅳ〕 標本分布の問題 (Problems of sampling distribution)

ところでこのうち〔Ⅱ〕と〔Ⅲ〕とについては，その数学的構成を述べた。〔Ⅰ〕は，前節の記号を用いれば，分布関数 $\mathrm{F}(x_1, x_2, \cdots, x_n)$ の関数型をいかに選ぶべきかという問題を意味すると，R. A. Fisher は規定している。たとえば，任意抽出標本であるとき各標本分布が正規分布に従うか否かというような問題は，正規分布を規定する２つの母数mとσとに関する仮説の検定あるいは推定

の問題と切離して，それよりもさきに論じておかなければならないというふうに解釈するのが，R. A. Fisher のいき方であり，したがって，〔II〕および〔III〕においては，ある有限個のパラメーターに関する仮説検定，あるいは推定の問題として取り扱おうというのが，R. A. Fisher 以来ネーマン－ピアソンのいき方であった。この研究方針に対しては，おそらく反省を要するものがあり，この分離の妥当性の検討は，統計推理の発展の鍵ともなろう。〔IV〕は，以上の問題を取り扱うための手段ともなるべきもので，内容的には，諸々の確率変数に関する四則，一次形式，一般にいうと，確率変数（$X_1, X_2, \cdots, X_n$）の分布関数を知って，それら確率変数の関数 $G(X_1, X_2, \cdots, X_n)$ の分布関数を求める問題であるといえる。$x_1, x_2, \cdots, x_n$ の関数 $G(x_1, x_2, \cdots, x_n)$ を**統計量**（statistic）というのであるが，この統計量に対する確率変数にほかならない。たとえば $x_1, x_2, \cdots, x_n$ なる標本に対して，標本平均値 $\bar{x}$，標本標準偏差 s，あるいは t はいずれも統計量である。ところでよく知られているように，$\bar{x}$ および s はそれぞれ母集団平均値 m および母集団標準偏差の**推定**（estimate）である。〔I〕,〔II〕および〔III〕を論ずるに当たって，われわれは，これらの統計量あるいはその特別の場合として推定量の確率分布を用いなければならない。この確率分布を求める問題が〔IV〕にほかならない。

ここでは，〔II〕および〔III〕に関して，重要な成果を紹介するにとどめよう。

A．**仮説検定に関するネーマン－ピアソンの理論**　この理論は，前節の用語を用いると，Ωが k 個の母数をもつ分布関数の集合にとった場合である。すなわち，母集団よりの標本抽出に関しては，未知ではあるが，ある一定の関数型をもつ分布関数を想定し，しかもそこにおける未知の母数はたかだか k 個とするのである。k 次元ユークリッド空間に，この k 個の母数の座標をもとめることにより，Ω の要素との間に一対一の対応をつけるのである。かくして Ω すなわち母数空間とみるのである。

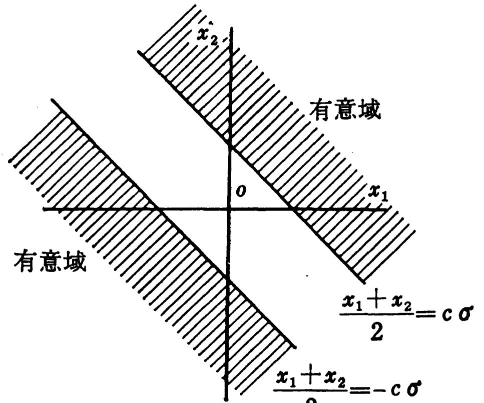
簡単な例として，任意標本抽出であって，標準偏差 σ は既知の正規分布を前

提し，平均値θのみを問題にする場合を考えよう。このとき

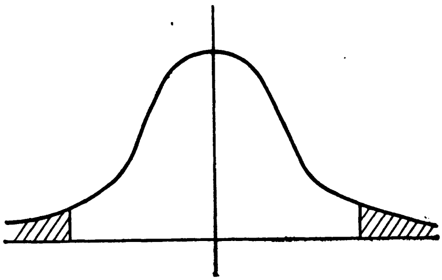
$$\text{(14)}\qquad F(x_1, x_2, \cdots, x_n ; \theta)=\frac{1}{(\sigma\sqrt{2\pi})_n}\prod_{k=1}^{n}\int_{-\infty}^{x_k} e^{-\frac{(u_k-\theta)^2}{2\sigma^2}}du_k$$

となっている。このとき $\theta=0$ かどうかを検定しようとする。

古典論においても，すでに記述統計学の項で述べたように，$|\bar{x}|\geqq c\sigma$であるとき，かつそのときにかぎり，$\theta=0$ という仮説は棄却していたのである。ここに常数 c の選び方は，次の原理によっていた。すなわち，$\theta=0$ という仮説のもとに $|\overline{X}|\geqq c\sigma$ ということの起こる確率が小さいときには，この仮説をすてる気になり得るように c をとる。このことは，ある小さな確率αをあたえて



11－2 図

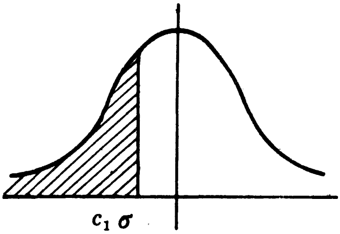


11－3 図

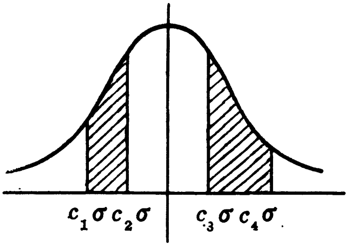
(15)　　$P(|\overline{X}|\geqq c\sigma)=\alpha$

これを満足するようなcをきめるのである。たとえば $\alpha=0.05$ とすれば $c=1.96$ である。

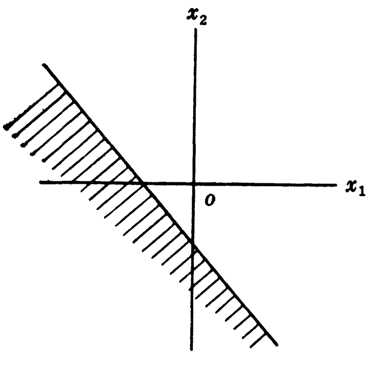
しかし，ここで重要なことは，古典理論にあっては，11-2 図のように，なぜこのような有意域を用いなければならないかの理由を明らかにしえなかった。上述の $\overline{X}$ は1つの統計域量に対する確率変数である。この論法は，次のことに同一である。$\overline{X}$ の分布関数を求めてみると，$\theta=0$ という仮説のもとにあっては，それは標準偏差 $\sigma/\sqrt{n}$ の正規分布である。上述の有意域の作り方は，この分布において，11-3 図のように，$|\overline{X}|$ がcより大きくなる「すそ」の方の面積を 0.05 にとってある。しかし，今まで述べた論法でいくと，何も「すそ」の方に対応に切り捨てる部分をつくらなくとも，ほかにいくらも



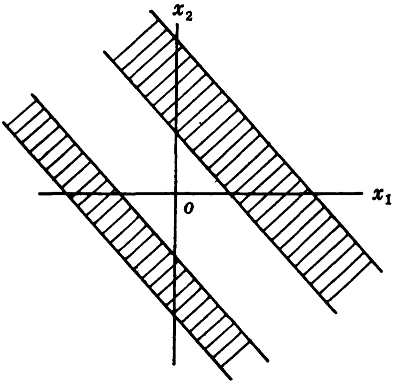
11-4 図



11-5 図



11-6 図



11-7 図

方法はある。それは 11-4〜7 図によってみられるとおりである。

とにかく次のようにとればよい。

(16)　　$P(\overline{X} \leqq c'\sigma)=0.05$

$$P(c_1\sigma \leqq \overline{X} \leqq c_2\sigma)+P(c_3\sigma \leqq \overline{X} \leqq c_4\sigma)=0.05$$

このように，有意域のとり方には，いろいろ可能であるから，われわれはこのうちいかなるものを選ぶべきかの基準をもたなければならない。この点を明確に意識して形式化した点に，ネーマン－ピアソンの理論の画期的な意義を認めなければならない。仮説が真なるにもかかわらずこれを偽りなりとして仮説を棄却する過誤を**第Ⅰ種の過誤**といい，仮説が偽なるにもかかわらず，これを真なりとして採択する過誤を**第Ⅱ種の過誤**という。標本空間においてつくられた有意域Rに，標本が落ちる確率を**有意域Rの大きさ**（size）という。上例ではRの大きさは $\alpha=0.05$ であった。有意域の大きさとは，第Ⅰ種過誤の起こる確率にほかならないのである。

ネーマン－ピアソンの理論の骨子は，第Ⅰ種過誤を一定の指定値 α 以下にしたもののうちで，第Ⅱ種過誤を最小にするような有意域Rを求めることにあるといえる。未知の母数が1個の場合についていえば，仮説 $\theta=\theta_0$ のときに，Rに落ちる確率を α にとる。このようなRのうちで θ が他の値 θ_1 のとき，Rに落ちる確率が最小になるようにRを選ぶということである。すなわち原則的には

(17)
$$\begin{cases} P(R/\theta_0) \equiv \int_R dF(x_1, x_2, \cdots, x_n ; \theta=\theta_0)=\alpha \\ P(R/\theta_1) \equiv \int_R dF(x_1, x_2, \cdots, x_n ; \theta=\theta_1)=\cancel{\mathrm{Min}} \quad \text{Max} \end{cases}$$

$P(R/\theta_1)$ を $\theta=\theta_1$ に関する有意域Rの**検定力**（Power）という。Rを一定にしておいて，検定力を θ の関数とみるとき，**検定力関数**（Power function）という。

上例にかえって述べよう。今有意域の大きさ $\alpha=0.05$ としよう。そうして次の二種類の有意域RおよびR′をとる。すなわちRは $|\bar{x}| \geqq c\sigma$ によりR′

は $\bar{x}\geqq c'\sigma$ により，標本空間内に規定される領域であるとする。$c=1.96/\sqrt{n}$, $c'=1.64/\sqrt{n}$ となることは正規分布関数表から容易にわかる。

(18)　$$P(R/\theta)=P(|\bar{X}|\geqq c\sigma/\theta)$$
$$=\frac{1}{\sqrt{2\pi}\frac{\sigma}{\sqrt{n}}}\int_{|u|\geqq c\sigma} e^{-\frac{n(u-\theta)^2}{2\sigma^2}}du$$
$$=\frac{1}{\sqrt{2\pi}}\left\{\int_{-\infty}^{-1.96+\sqrt{n}\frac{\theta}{\sigma}}+\int_{1.96+\sqrt{n}\frac{\theta}{\sigma}}^{\infty}\right\}e^{-\frac{t^2}{2}}dt$$

となり，同様にして

(19)　$$P(R'/\theta)=P(\bar{X}\geqq c'\sigma/\theta)$$
$$=\frac{1}{\sqrt{2\pi}}\int_{1.64+\sqrt{n}\frac{\theta}{\sigma}}^{\infty}e^{-\frac{t^2}{2}}dt$$

そこで，これを図示するのには，$\sqrt{n}\theta/\sigma$ を便宜上 1 つの母数φとしてとることにして，

(20)　$$P(R/\theta)=P(R/\varphi)=\frac{1}{\sqrt{2\pi}}\left\{\int_{\infty}^{-1.96+\varphi}+\int_{1.96+\varphi}^{\infty}\right\}e^{-\frac{t^2}{2}}dt\equiv P_1(\varphi)$$
$$P(R'/\theta)=P(R'/\varphi)=\frac{1}{\sqrt{2\pi}}\int_{1.64+\varphi}^{\infty}e^{-\frac{t^2}{2}}d\varphi\equiv P_2(\varphi)$$

一般に同じ大きさの 2 つの有意域 R および R′ が存在するとき，θ のある値 $\theta=\theta_1$ において，R′の検定力曲線が R の検定力曲線の上にあるならば，θ の真値が θ_0 である場合に，第II種の過誤をおかす確率は，R′ の方が R よりもより小になる。したがって，R と R′ は第 I 種過誤をおかす確率は等しいが，第II種過誤をおかす確率は R′ の方がより小である。$\theta=\theta_1$ に対して R′ が R よりも**より有力** (more powerful) であるという。$\theta=\theta_0$ 以外のすべての θ の値に対して，R′ に対する検定力曲線がRのそれよりも上にあるならば，**一様に** R′ はRよりも**より有力** (uniformly more powerful) というべきものとなる。もしそういう事情が現出したとすると，Rをえらぶ理由は，一応ないので，R′

を当然選ぶべきである。もしも幸いにして，Rの検定力曲線が同じ大きさをもつほかのいかなる有意域 R′ の検定力曲線の下になることがないならば，Rは**一様に最有力** (uniformly most powerful) といわれる。しかし多くの場合，一様に最有力な有意域というものはないのである。

上例についてみても，R と R′ とは，一方が他方より一様にはより有力となっていない。しかし，次のことは見得られるであろう。θ は必ず負にはならないということが a priori に知られているならば，$\theta \geqq 0$ だけを考えればよいから R′ は R よりも一様により有力であるといえるし，同様にして，$\theta \leqq 0$だけを考えてよい場合には，標本空間において $x \leqq c'\sigma$ で定義される有意域 R″ の方が，R あるいは R′ よりも，一様により有力であるということはいえる。このように当の資料以外の知識が充分に活用されなければならない。とくにこれらの知識がなければ R は R′ と R″ との何よりも，より妥当であると考えられる。それらがもっと一般に考えて次の点からも考えられよう。

このことは，ネーマン－ピアソンの導入した原理として**不偏性の原理**ということを，有意域選定の基準条件にとることによって明らかになるであろう。ある検定法，すなわちある有意域が不偏であるというのは，その検定力関数が，検定しようとする母数値 $\theta=\theta_0$ の付近において局部的に極小になっていることである。この原理を基礎づけるのには，もしかりに，この不偏性原理が満足されてない有意域を考える（たとえば上例のR′)。 すると，仮説を却棄する確率が $\theta=\theta_0$ のときよりも，θ_0 でないある θ_1 のときにおいて，より大きいということになる。これは望ましい状態ではないといわなければならない。ところでこの不偏性を満足し，しかも同一の大きさをもつ有意域をもつ諸検定法のうちで，他の対立仮説のすべてに関して他のいかなる検定法にもまさるとも劣ることのない検定法，いえかえれば少なくとも同等またはより以上に有力な有意域をもつ検定法を，**一様最有力不偏検定** (uniformly most powerful unbiassed test) という。このような検定法に対応する有意域をネーマン－ピアソンは**A_1型有意域**と称したのである。

上例についていえば，R は A_1 型有意域であることが証明できる。A_1 型有意域は，重要な場合に見いだされるのではあるけれども，他方，多くの場合にそういうものが存在しないということもいえるのであって，ネーマン－ピアソンがさらに進んで第3の型を導入したのが，すなわち**A型有意域**というものであって，ある有意域RがA型であるというのは，それの検定力関数が次の性質を有することである。

$$(21)\qquad (1°)\quad \left(\frac{\partial P(R/\theta)}{\partial\theta}\right)_{\theta=\theta_0}=0$$

$$(2°)\quad \left(\frac{\partial^2 P(R/\theta)}{\partial\theta^2}\right)_{\theta=\theta_0} \geqq \left(\frac{\partial^2 P(R'/\theta)}{\partial\theta^2}\right)_{\theta=\theta_0}$$

ここに R' は (1°) の条件を満足し，R と同じ大きさをもつ任意の有意域である。

条件 (2°) は，簡単にいえば，θ_0 の付近では R が最も有力であると主張するものと，だいたいみてよいであろう。A型有意域は，多くの場合に存在し，実用上重宝である。ただ，これは検定しようとする値 $\theta=\theta_0$ の付近だけの検定力の比較で最有力というだけであるが，実際問題となるのは相当離れた θ の値である，という非難もあったわけである。しかしこれらの非難は最近の研究によれば，かならずしも当たらないようである。

B.　**R. A. Fisher の推定論**　標本 $x_1, x_2, \cdots, x_n$ の適当な関数 $T_n=t_n(x_1, x_2, \cdots, x_n)$ を選び，これをもって母集団の未知母数 θ の推定値としようとするとき，いかなる関数をとるのがよいか，あるいは最良かという問題がおこる。われわれの立場は，$x_1, x_2, \cdots, x_n$ をもって確率変数 $X_1, X_2, \cdots, X_n$ の実現値とみるのであるから，問題は，確率変数 $t(X_1, X_2, \cdots X_n)$ を規定する関数 t の決定にあるといわなければならない。すなわち，関数 t を適当にとることによって，未知母数の真の値の近くにある確率をできるだけ大きくすることであると，概括していうことができよう。

推定の方法とは，この場合関数 t の選定法に帰着するのであるが，これに関

して，まず第1に常識的にあげられるであろうことは，標本の大きさを限りなく大きくして全集団にいたるようになるならば，その推定は真の値をあたえなければならないということであろう。数学的にいえば，T_n がなんらかの意味において θ へ収束しなければならない。この条件を満足する統計量を**一致統計量** (consistent statistic) という。R. A. Fisher の推定論の出るまで，Markov 等の研究を別とすると，この一致性のみが推定法選定のただ一つの基準であった。しかし，ここに多くの問題がある。

マルコフの基準というのは次のようなものである。Markov は次の条件を満足するとき，この統計量を最良推定 (best estimate) といった。

(1°)　第1に確率変数 $t(X_1, X_2, \cdots, X_n)$ は不偏である。すなわち

(22)　　$$E_0\{t(X_1, X_2, \cdots, X_n)\}=\theta$$

ここに左辺は，未知母数の真の値が θ あるという条件のもとにおける，確率変数 $T=t(X_1, X_2, \cdots, X_n)$ の平均値を意味する。

(2°)　上の条件 (1°) を満足するあらゆる統計量のうちで，とくに t は他の t' に対して次の関係を満足する。

(23)　　$$E_0\{(t(X_1, X_2, \cdots, X_n)-\theta)^2\} \leqq E_0\{(t'(X_1, X_2, \cdots, X_n)-\theta)^2\}$$

この基準は一応は合理的である。チェビチェフの定理を用いるならば，標準偏差が小さいこと，θ の付近に統計量の値がより多く集中するということは，一応緊密に相関することであるのはわかる。しかしながら，これは，絶対的にいえることでないから，標準偏差だけで片づけようとするのは無理である。また，(1°) および (2°) を満足するような推定は多くの場合存在しないのである。

しかし，これは，標本の大きさ n に関してなんらの制限も設けなかった場合の議論であった。とくに n を限りなく大きくするということであれば，事情はかえって簡単になる。

たとえば，正規分布の母集団の任意標本から，その母集団標準偏差 σ を推定するに当たって，次のいずれかの推定を用いることが多い。

$$\text{(24)}\qquad \sigma_1(\mathrm{E}_1)=\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum|x_i-x| \quad \text{（標本平均誤差）}$$

$$\sigma_2(\mathrm{E}_2)=\sqrt{\frac{1}{n}\sum(x_i-\bar{x})^2} \quad \text{（標本標準誤差）}$$

これらは，いずれも一致統計量である。したがって，一致性だけを推定法の基準にとれば，計算のより容易な標本平均誤差を用いるのが当然である。ところが，他の基準を入れて考えると両者には大きな相違がある。

R. A. Fisher は，**有効統計量** (efficient statistic) の概念を導入した。すなわち n が大きくなるにつれ，平均値 θ の正規分布に近づく一致統計量のうちで，最小の標準偏差を有するものを意味するのである。場合によっては，そういうものは存在しないかもしれない。また幾通りも存在し得るかもしれない。とにかくこのように平均値 θ の正規分布に近づく統計量どうしにおいては，標準偏差が n とともにどれほど小さくなるかが問題なのであって，有効統計量の標準偏差との比較において，**統計量の効率**(efficiency)というものが導入される。

少し厳密に表現しよう。次の条件を満足するとき統計量の系列 $\{t_n\}$ に対応する確率変数列 $\{\mathrm{T}_n\}$ は有効であるという。

(1°)　$\sqrt{n}(\mathrm{T}_n-\theta)$ は，n が限りなく大きくなるとき，平均値 0，有限な標準偏差 σ をもつ正規分布に近づく。

(2°)　条件 (1°) を満足する他のいかなる $\{\mathrm{T}'_n\}$ に関しても，次の関係をみたす。

$$\text{(25)}\qquad \sigma^2/\sigma'^2 \leqq 1$$

ただし

$$\text{(26)}\qquad \sigma^2=\lim_{n\to\infty}\mathrm{E}_0[(\sqrt{n}(\mathrm{T}_n-\theta))^2]$$

$$\sigma'^2=\lim_{n\to\infty}\mathrm{E}_0[(\sqrt{n}(\mathrm{T}_n'-\theta))^2]$$

σ^2/σ'^2 を $\{\mathrm{T}_n'\}$ の効率という。

それには，標本の大きさ n が大きくなるとき，一致統計量 T_n が推定しようとする θ へ近づく模様を問題にする。もちろんそれは確率論的な意味において

であるが，θ への収束する速さというものが問題になるのである。たとえば T_n の分布状態を考える。多くの実際上の場合において，T_n の関数型からして，大数の法則が適用されるから，一致統計量の分布関数は，平均値 θ で，標準偏差が n とともに小さくなる正規分布関数に近づくのである。そういうときには，標準偏差が n とともに早く 0 になるものの方が，θ を推定する統計量として有効であるといえる。

たとえば，標準偏差 σ が既知で，平均値 m が未知の正規母集団からの任意標本抽出にあって，m を推定するのに，

$$(27) \qquad t_n=\frac{x_1+x_2+\cdots+x_n}{n}$$

$$(28) \qquad t_n'=(t_1, t_2, \cdots, t_n \text{ の中央値})$$

というふうに統計量をとる。すると両者はともに一致統計量ではあるが

$$\lim_{n\to\infty}\sigma^2(T_n)=\sigma^2, \quad \lim_{n\to\infty}\sigma^2(T_n')=\frac{\pi}{2}\sigma^2$$

それゆえに，大きさ 100 の標本から求めた標本平均値を用いることと，大きさ $100\times\pi/2=157$ の標本から求めた標本の中央値を用いることが同程度の精確さをもつことになる。標本が与えられたとき，中央値で母集団平均値を推定しようとするのは，標本平均値の場合に比べると，全標本の $2/\pi$ だけしか使ってないという意味にもなる。しかしながら，もし母集団分布がラプラスの分布 $e^{-|x-m|}/2$ になると，今度は目標の計算で事情は反対に，$\sigma^2(T_n')/\sigma^2(T_n)$ の比は 1/2になるから，大きさ 100 の標本よりの標本平均値は，大きさ 50 の標本からの標本中央値にしか精確さが当たらないということになる。

以上の一致統計量および有効統計量という概念は，標本の大きさ n が充分大きな場合に使用されるべき概念である。n がそれほど大きくない場合に対しても，統計量の適否を規定する概念があって欲しい。R. A. Fisher は，このために**本源的精確度**（intrinsic accuracy）という概念を導入した。これは標本の大きさが何であっても定義できるものである。本源的精確度は必ず 0 より小

さくはなく1より大きくもないものであって，その値が1に近いほど，統計量としての精確度は高い。とくにこれが1に等しいときには**充足統計量**（sufficient statistic）といわれる。充足統計量 $t(x_1, x_2, \cdots, x_n)$ というのは次のような性質をもつものであると定義してもよい。すなわち

(29)　$$t(x_1, x_2, \cdots, x_n)=t(x_1', x_2', \cdots, x_n')$$

となる場合には

(30)　$$\frac{\mathrm{F}(x_1, x_2, \cdots, x_n/\theta)}{\mathrm{F}(x_1', x_2', \cdots, x_n'/\theta)}$$

は θ に無関係である。すなわち $\mathrm{F}(x_1, x_2, \cdots, x_n/\theta)=\mathrm{G}(t/\theta)\mathrm{H}(t)$ という形で書けることである。このことを換言すれば，$t(x_1, x_2, \cdots, x_n)$ は，次の条件を満足するとき，θ を推定するための充足統計量であるといえる。すなわち他の任意の統計量 $t'(x_1, x_2, \cdots, x_n)$ に対して，これらに対応する確率変数 $t(\mathrm{X}_1, \mathrm{X}_2, \cdots, \mathrm{X}_n)$ と $t'(\mathrm{X}_1, \mathrm{X}_2, \cdots, \mathrm{X}_n)$ とに関して次のことが成り立つことである。すなわちそれは

(31)　$$\mathrm{P}\{s<t'(\mathrm{X}_1, \mathrm{X}_2, \cdots, \mathrm{X}_n)<s+ds/t(\mathrm{X}_1, \mathrm{X}_2, \cdots, \mathrm{X}_n)=t\}=f(s/t)\,ds$$

と書くとき $f(s/t)$ が θ に無関係であるということである。すなわち $t(x_1, x_2, \cdots, x_n)$ を知れば，θ を推定するのに $t'(x_1, x_2, \cdots, x_n)$ を追加しても，さらに知識はふえはしないという意味になる。θ を推定するために $(x_1, x_2, \cdots, x_n)$ からくみとれる知識はことごとく $t(x_1, x_2, \cdots, x_n)$ のなかにはいっている（exhaust information）という意味にもなるのである。

たとえば，標本 $(x_1, x_2, \cdots, x_n)$ が平均値 m，標準偏差 σ の正規母集団からの任意抽出標本であるとすると，$\mathrm{F}(x_1, x_2, \cdots, x_n/m, \sigma)$ に対する確率密度は

(32)　$$f(x_1, x_2, \cdots, x_n/m, \sigma)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-m)^2}{2\sigma^2}}$$

となり，しかもこれは右辺をかきかえることによって

(33)　$$(2\pi\sigma^2)^{-\frac{n}{2}}e^{-\frac{1}{2}(\bar{x}-m)^2/\sigma^2}e^{-\frac{S}{2\sigma^2}}$$

となる。ただしここに $\bar{x}$ は標本平均値，S は平方和であって

$$\text{(34)} \qquad S=\sum_{i=1}^{n}(x_i-\bar{x})^2$$

式から，統計量 $\bar{x}$ は m に対しての充足統計量であることがわかる。

以上，推定の良否を判定する基準ともいうべきものとして，一致性，有効性および充足性を導入したのは R. A. Fisher の業績である。この 3 つの性質をあわせもった統計量を，**最適統計量** (optimum statistic) という。Fisher はこの最適統計量を求めることを目標とした。これを見いだすためには R. A. Fisher は**最尤法の原理** (principle of maximum likelihood) という方法を導入したのである。これを利用するためには $(X_1, X_2, \cdots, X_n)$ に対応する確率分布 $F(x_1, x_2, \cdots, x_n/\theta)$ の微分可能性を仮定する。すなわち

$$\text{(35)} \qquad F(x_1, x_2, \cdots, x_n/\theta)=\int_{-\infty}^{x_n}\int_{-\infty}^{x_{n-1}}\cdots\int_{-\infty}^{x_1}f(u_1, u_2,\cdots, u_n/\theta)\,du_1, du_2,\cdots du_m$$

と書けるものとする。このとき標本点 $(x_1, x_2, \cdots, x_n)$ に対する尤度(likelihood)というのは，この点に対する確率密度のことであって

$$\text{(36)} \qquad L=f(x_1, x_2, \cdots, x_n/\theta)$$

としてあらわされるものを意味する。L は θ の関数でもあるが，この L を最大にするような θ の値を $\theta_n(x_1, x_2, \cdots, x_n)$ とするとき，この値をもって，θ の推定とするのを，最尤法による推定というのである。R. A. Fisher の推定論における主要結果は，次のようにいいあらわせるであろう。

$\{x_n\}$ を同一母集団からの（返却的な）任意抽出標本であるとすれば，この母集団の従う分布関数が若干の条件を満足するとき最尤法推定 $\theta_n(x_1, x_2, \cdots, x_n)$ は前述の意味における有効統計量となる。

充足統計量をもつための母集団分布関数を特徴づける問題は，R. A. Fisher, Hotelling, Doob 等によって完成されている。上述の最尤法解は，充分大なる n に対しては，近似的に最適統計量となる。

C. **Neyman の信頼区間の理論** 上述の R. A. Fisher の理論では母集団

の母数は未知ではあるが，ある確立したものであると前提して論じた。ここでも，この前提を設けることという点においては，なんら変わりはない。ところで R. A. Fisher の前述の理論では，とにかくこの未知の母数を，これこれの値であるとしてある1つの値として，なんらかの方法で推定しようとするものである。これに対して，この未知の母数は，たいていの場合は，これこれの範囲内にあるとみてよいという区間を指示する推定方法が Neyman によって展開されている。前者が**点推定**（point estimation）といわれるのに対して，後者は**区間推定**（interval estimation）といわれるものである。後者の方がより実際的であることはいうまでもない。

区間推定法にあっては，標本点 $e=(x_1, x_2, \cdots, x_n)$ の関数として，$\bar{\theta}(e)$ および $\underline{\theta}(e)$ なる2つの関数を導入して，真値 θ を区間 $\delta(e)=(\underline{\theta}(e), \bar{\theta}(e))$ のなかに含むようにすることが問題なのである。

ところで，標本点 e に対して確率変数 $\mathrm{E}=(\mathrm{X}_1, \mathrm{X}_2, \cdots, \mathrm{X}_n)$ を対応させて考えると，$\underline{\theta}(\mathrm{E}), \bar{\theta}(\mathrm{E})$ もまた確率変数であるからして，当然

(37)　　$\underline{\theta}(\mathrm{E}) \leqq \theta \leqq \bar{\theta}(\mathrm{E})$

という事象の確率ということは考えられる。

$\underline{\theta}(\mathrm{E}), \bar{\theta}(\mathrm{E})$ なる2つの確率変数は次の条件を満足するとき未知母数 θ の**信頼区間**（confidence interval）であるという。

(1°)　すべての $\mathrm{E}=(\mathrm{X}_1, \mathrm{X}_2, \cdots, \mathrm{X}_n)$ に対して $\underline{\theta}(\mathrm{E}) \leqq \bar{\theta}(\mathrm{E})$

(2°)　θ のすべての値に対して，次の等式が成り立つ。

(38)　　$\mathrm{P}(\underline{\theta}(\mathrm{E}) \leqq \theta \leqq \bar{\theta}(\mathrm{E})/\theta)=\alpha$

ここに α は一定数である。左辺は未知母数が θ であるとき，この θ が，$\underline{\theta}(\mathrm{E}) \leqq \theta \leqq \bar{\theta}(\mathrm{E})$ を満足する確率をあらわす。α は**信頼係数**（confidence coefficient）といわれる。

信頼区間の実際的な意義は，つぎのようである。大きさ n の標本を，幾組も多数回にわたってつくったとしよう。標本の各組に対して，未知数 θ は区間 $[\underline{\theta}(\mathrm{E}), \bar{\theta}(\mathrm{E})]$ に属するという判断を下すならば，ある時は正しくて成功し，

ある時は誤りであって失敗するということになるから，その成功率はαに等しいということになる。

一般にいうと，αが大であればあるほど，信頼区間は長くなり，αが小になればなるほど信頼区間は短くなる。信頼区間の短いということはつまり鋭い推定であるということであり，信頼係数の大きいことは信用のおける推定であるということになる。鋭いということと，信頼のおけるということは当然，いわば相補的なものであることをみるであろう。信頼区間の選び方の問題というのは，信頼係数の方は指定しておいての議論が多い。

信頼係数αが与えられたとき，上述の信頼区間の選び方は一般に無限に多くあるわけであるが，一般的にいうと，この区間の長さが短いほどよいと考えられる。このためには，最短信頼区間と，不偏信頼区間等の諸概念を導入するのであるが，最短信頼区間というのは，次のような信頼区間 $(\underline{\theta}(\mathrm{E}), \bar{\theta}(\mathrm{E}))$ を意味する。すなわち上述の (1°)，(2°) を満足する他のいかなる信頼区間$(\underline{\theta}'(\mathrm{E}), \bar{\theta}'(\mathrm{E}))$ と，母数のいかなる値 θ' と θ'' とに対しても

$$(39)\qquad P(\underline{\theta}(\mathrm{E}) \leqq \theta \leqq \bar{\theta}(\mathrm{E}) \mid \theta'') \leqq P(\underline{\theta}'(\mathrm{E}) \leqq \theta \leqq \bar{\theta}'(\mathrm{E}) \mid \theta'')$$

このような最短信頼区間がもし存在すれば，推定としては最も有利なものである。しかし実際には存在するのがむしろ例外的なことで，一般には存在しない。このために Neyman の上述の仮説検定法と同じく不偏性原則 (principle of unbiasedness) を導入する。

信頼区間 $\delta(\mathrm{E})$ が上述の (1°)，(2°) のほかに，なお次の条件を満足するとき，これを信頼係数 α に対する**不偏信頼区間** (unbiasedness confidence interval) であるという。すなわちそれは母数のすべての値 θ', θ'' に対して，

$$(40)\qquad P(\underline{\theta}(\mathrm{E}) \leqq \theta' \leqq \bar{\theta}(\mathrm{E}) \mid \theta'') \leqq \alpha$$

という条件を意味する。同じ信頼係数αに対する不偏信頼区間のうちで，上述の意味で最短信頼区間でもあるものを**最短不偏信頼区間** (shortest unbiased confidence interval) という。もしこういうものが存在すれば，これは，不偏性原則という資格基準を必要とするかぎり，最も有利な推定といわなければな

らないわけであるが，不幸にして，これもごく制限された場合にしか存在しないものである。このために Neyman は，short unbiased confidence interval という概念を導入した。これは**局所最短不偏信頼区間**とも訳すべきものであって，一定の信頼係数に対する不偏信頼区間のなかで，いかなる θ' に対しても，また他の同様な性質をもついかなる $(\underline{\theta}'(\mathrm{E}), \bar{\theta}'(\mathrm{E}))$ に対しても

$$(41)\quad \left[\frac{\partial^2}{\partial\theta''^2}\mathrm{P}(\underline{\theta}(\mathrm{E})\leqq\theta'\leqq\bar{\theta}(\mathrm{E})\,|\,\theta'')\right]_{\theta''=\theta'}\leqq\left[\frac{\partial^2}{\partial\theta''^2}\mathrm{P}(\underline{\theta}'(\mathrm{E})\leqq\theta'\leqq\theta'(\mathrm{E})\,|\,\theta'')\right]_{\theta''=\theta'}$$

を満足するものを意味する。

Fisher にしても，Neyman にしても，これらの推定理論においてベイズの定理が用いられていないことに注目すべきであろう。ことに Fisher の理論は常にベイズの定理の排撃ということを1つの目標においてあったことは，すでに述べたとおりであった。ベイズの定理の正しい使用法にあっては，推定しようとする θ は未知ではあるが，とにかく一定の値をもつという R. A. Fisher の前提とはことなり，θ はそれ自身1つの確率変数 Θ の現実値とみなければならないのである。この Θ が $(\theta, \theta+d\theta)$ の間にある確率を $g(\theta)d\theta$ とし，$\Theta=\theta$ のときの標本に対応する確率変数Eの分布密度を $f(x_1, x_2, \cdots, x_n\,|\,\theta)$ とすれば，ベイズの定理により $\mathrm{E}=(\mathrm{X}_1, \mathrm{X}_2, \cdots, \mathrm{X}_n)$ の実現値 $e=(x_1, x_2, \cdots, x_n)$ を知って，Θ が $(\theta, \theta+d\theta)$ の間にあった確率は

$$(42)\qquad \frac{g(\theta)f(x_1, x_2, \cdots, x_n/\theta)}{\int_{-\infty}^{\infty} g(\theta)f(x_1, x_2, \cdots, x_n/\theta)d\theta}$$

で表わされる。

しかしこの考え方には2つの欠点がある。第1に真の値 θ というものを確率変数の実現値とみることが，一般にはかならずしも妥当でない。これを認めるには，母集団の母集団ともいうべきものを導入しなければならない。もちろんそういう必要のある場合もあり，確かにそう考えられる場合もあるけれども，一般には，これは無用の前提をつけ加えるものである。第2に，またかりに，第1の点は容認するとしても，関数 $g(\theta)$ が一般には不明であるという欠点が

ある。漠然とした根拠からこれに種々の関数を想定することは正確な統計推理のためにとらないというのが, R. A. Fisher の意見である。たとえば事前確率といわれる $g(\theta)$ について, 無知を理由として一様分布を仮定する。すなわち, $a \leqq \theta \leqq b$ において $g(\theta)=$一定 とすればよいようである。しかし, Θ の分布が一様分布であると考えられると同様に, Θ の他の関数, たとえば Θ^2, Θ^3 等もまた一様分布であるとも考えられる。そうして, Θ が一様分布であることと, たとえば Θ^2 が一様分布であることとは矛盾するのである。

ベイズの定理の排撃という点では Fisher が最も徹底的であった。Fisher は現実の資料以外の知識, これを a priori な知識というが, それに依存しないことを目標にした理論構成を理想にしている。しかしながらベイズの定理を用いることは, 上述のような欠点はあるけれども, さらに仔細に考えるならば, なお再検討の余地はあることをつけ加えておきたい。

D.　**A. Wald の推定論**　推定理論の顕著な1つの進展は 1940—1941 年の間において A. Wald によって与えられた。

上述のように一様最有力（不偏）検定（uniformly most powerful <unbiased> test）もこれに対応する最短（不偏）信頼区間（shortest <unbiased> confidence interval）も, いずれも概念としては適切なものであり, これにより一応最良の検定, 最良の推定が得られるわけではあるけれども, 遺憾ながら, それの存在するのはごく限られた特殊の場合にすぎないのである。これに代わるものとして Neyman のいわゆるA型検定や最短信頼区間もあるけれども, これらは, 問題の未知母数の値として特殊の値の付近だけしか考慮していない。ところが実際には, 問題の値の近くだけでなく, むしろ相当離れた値との比較こそ問題なのであるから, 基準を局所的に解釈する点に実際上不便があると考えられていた。ところが上述の研究によって, A. Wald は, 以上の困難は, いわば外見的なものであることと, 推定理論としては, もっと突っ込んで主張できることを示したのである。一様最有力検定, 最短信頼区間は一般の確率分布に対しては存在しないのであるけれども, これらは, 存在しなくてさ

しつかえない，なぜなら，これに代わるものとして，漸近的最有力不偏検定（asymptotically unbiased most powerful tests）と漸近的最短不偏信頼区間（asymptotically shortest unbiased confidence interval）とは，nが充分大きくなるにつれ，それぞれこれらの代わりになるものであり，しかもこれらは，実際上ではほとんどすべての場合について存在するといえることを，A. Wald は示したのである。しかし，本書においてはその内容には立ち入らないことにする。

7.　推測統計学の最近の進歩

この方面の最近の進歩は，ネーマン－ピアソンの到達した段階をこえて進出しつつある。今この方面に最も貢献しつつある A. Wald による On the Principles of Statistical Inference（1943年）の所説を紹介しよう。

A. Wald の指摘するように，Fisher およびネーマン－ピアソンの理論は2つの点において制限されたものである。第1に，彼らの取り扱った問題が，仮説検定論と，推定論という2つの問題に局限されたことである。第2に，問題の対象たる分布関数の世界を，ある有限次元の母数で表現される分布関数の集合に局限したことである。

事実，実際上必要な問題であって，しかもこれらの理論において取り扱い得なかったものとして，A. Wald は次のような例をあげている。

（1°）回帰曲線として多項式を採用してよいということは前提されているとしよう。このとき問題は資料に当てはめるべき多項式の次数いかんということである。今 H_n をもって関数関係がn次多項式であらわされる母集団からの標本であるという仮説をあらわす。このとき問題は，資料をもとにして判断した場合，$\mathrm{H}_1, \mathrm{H}_2, \cdots, \mathrm{H}_n, \cdots$ の仮説のうちいずれをとるかを決定するかということである。この問題は，ネーマン－ピアソンの仮説検定論にはいらないと，A. Wald はいうのである。

（2°）k 個の母数をもって表現される分布関数の集合としてΩが解釈されな