Estimationの問題について

古屋, 茂
中央航空研究所

出版情報：統計数理研究. 1 (2), pp.12-32, 1942-03-15. Research Association of Statistical Sciences

Estimation の問題について

會員 古屋 茂 (中央航研)

(昭和十六年七月十八日受理)

Estimation の問題は統計數學に於ける重要な問題の一つであつて 1920年以後發達して來た．その理論は主として R.A.Fisher に負ふものである．ここでその大要を紹介しよう．この理論の出發點は次の問題である．

問題：p ケのパラメーター $m_1, m_2, \cdots, m_p$ を含む確率密度分布 $f(x; m_1, m_2, \cdots, m_p)$ が與へられたとする．この分布に關して $n(\geqq p)$ 回の觀測を行つて，$(x_1, x_2, \cdots, x_n)$ なる一組の値——之を size n の sample といふ——を得たとき，この觀測結果より未知常數 $m_1, \cdots, m_p$ を如何に定めるべきか？

$(m_1, \cdots, m_p)$ を觀測結果 $(x_1, \cdots, x_n)$ の函數として定める方法は種々あり，それ等は各々の立場に於て夫々の意味をもつものであるが，以下に於て確率論的にこの問題が如何に取扱はれるか，又未知常數を如何に定めるのが妥當であるかを說明しよう．

先づ注意すべきことは，$m_1, \cdots, m_p$ は各々一つの値——それが未知であるが——をとるのであつて，種々の値を夫々の確率でとるのではないことである．從つて所謂 a posteriori の確率を與へる Bayes の定理をそのまゝ使用することは出來ない．

以下 I に於てはパラメーターの値を一ケの値によつて推定する方法を述べ，II に於ては區間によつてパラメーターの値を推定する方法を述べる．

I. 一ケの値による推定法

sample $(x_1, \cdots, x_n)$ に於て各 x_i は，その確率密度分布として $f(x; m_1, \cdots, m_p)$ を有する確率變數 (240頁, 定義 2.1.)* と考へられるが，當分は n ケの確率變數 $x_1, \cdots, x_n$ が相互に獨立 (247頁, § 1.2.) であると假定する．更に，特に斷らない限り集合 $(x; f(x; m_1, \cdots, m_p) > 0)$ ——之を f の range といふ——は $(m_1, \cdots, m_p)$ に無關係なものと假定する．

§ 1. Method of Moments, Statistic, Consistent Statistic

上記問題の一つの解法としてよく使用されるものに method of moments がある．それは分布自身の moments と sample moments とを等しいとおいてえられる p ケの等式

$$(1.01) \qquad \frac{1}{n}\sum_{i=1}^{n} x_i^k = \int x^k f(x; m_1, m_2, \cdots, m_p)\,dx \quad (k=1, 2, \cdots, p)$$

より $m_1, m_2, \cdots, m_p$ を定める方法である．これは n が大きければ大きい程正確な方法といふ事が出來よう．その理由は

$$(1.02) \qquad E(X^k) = \int x^k f(x; m_1, m_2, \cdots, m_p)\,dx$$

が存在するならば，$n \to \infty$ のとき (1.01) の左邊は $E(X^k)$ に法則收斂，及び概收斂 (法則收斂，概收斂については 244頁, § 4.1 參照) するからである．この方法の長所は計算の實行が容易な點にあ

る．K. Pearson は通常よく現はれる確率密度分布を 7 種の型に分類したが，その際にこの方法を使用した．又最初の問題で「p ケのパラメーターを含む」といふ代りに「可附番ケのパラメーターを含む」といふ事にすれば，estimation の問題といふよりもむしろ curve fitting の問題となるが，その場合にも method of moments が便利である．例へば

$$\frac{d^n}{dx^n}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\right)=\sqrt{n!}\,\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}p_n(x) \tag{1.03}$$

から定まる多項式 $p_n(x)$ に依つて展開される確率密度分布 $f(x)$ を考へる：

$$f(x;a_1,a_2,\cdots,a_n,\cdots)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}(1+a_1p_1(x)+a_2p(x)+\cdots+a_np_n(x)+\cdots) \tag{1.04}$$

然るに

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}p_n(x)p_m(x)\,dx=\delta_{m,n} \tag{1.05}$$

なる故

$$E(p_i(x))=a_i\,. \tag{1.06}$$

従つて sample $(x_1,x_2,\cdots,x_n)$ より $\sum_{j=1}^{n}p_i(x_j)/n$ を作ればそれは a_i に概收斂するのである．

次に $E(X^k)$ の存在が保證されて居ない場合はどうであらうか．今 $p=1$ として

$$\sum_{i=1}^{n}x_i/n \tag{1.07}$$

が常數 a に法則收斂する爲の必要且十分な條件を求めよう．以下，(1.07) 即ち sample mean を $\bar{x}$ で表はすことにする．

$f(x;m)$ の特性函數 (241頁，§3.2) を $\varphi(t,m)$ とおけば，確率變數 (1.06) の分布の特性函數は $\varphi^n(\frac{t}{n},m)$. (251頁，定理 2.5.) 従つて $\psi(t,m)=\log\varphi(t,m)$ とおけば求める條件は，$n\to\infty$ の時 $n\psi(\frac{t}{n},m)\to ita$ である．(251頁，定理 2.5) 換言すれば $\psi(t,m)$ が $t=0$ で微分可能でその微係數が ia に等しいことである．平均値 $E(x)$ が存在するときは上記 $\psi(t,m)$ は $t=0$ で微分可能で，その微係數は $iE(x)$ に等しい．この逆が必ずしも成立しないこと――$\psi(t)$ の原點に於ける微分可能性から $E(x)$ の存在は出ないこと――は例へば $a<1$ として

$$f(x)=\begin{cases}\dfrac{K}{x^2(\log|x|)^a} & [\,|x|>2\text{ のとき}\,]\\ 0 & [\,|x|\leqq 2\text{ のとき}\,]\end{cases} \tag{1.08}$$

――但し茲に

$$\frac{1}{K}=\int_{|x|>2}\frac{dx}{x^2(\log|x|)^a} \tag{1.09}$$

とする――により容易に分る．

結局 $\psi(t,m)$ が $t=0$ で微分不可能の場合には $\bar{x}$ を作つても何等の効果のない事が分つた．

例として Cauchy の分布

$$f(x,m)=\frac{1}{\pi\{1+(x-m)^2\}} \tag{1.10}$$

をとる．その特性函数は $e^{-|t|+itm}$ なる故，$\bar{x}$ の分布の特性函数も亦 $e^{-|t|+itm}$．よつて $\bar{x}$ の分布も亦 $f(x, m)$ となる．それでこの場合 m を推定する爲に sample の平均値をとる事は $n=2$ としても $n=100$ としても更に n が何であつても全く同様に無意味な事である．method of moments は以上の如き缺點を持つてゐる．しかしその他にも種々の方法がある．

定義 1. sample $(x_1, x_2, \cdots, x_n)$ の一價可測函數 $T_n(x_1, \cdots, x_n)$ が或パラメーターの値 $g(m)$ を推定するものとして使用される時，T_n を statistic，又は $g(m)$ の estimate といふ．T_n は確率變數であるが $n\to\infty$ のとき $g(m)$ に法則収斂するならば，T_n を consistent statistic，又は $g(m)$ の consistent estimate といふ．statistic T_n の平均値 $E(T_n)$ が $g(m)$ ならば，T_n を $g(m)$ の unbiassed estimate と呼ぶ．

次に二三の例，$g(m)$, $\sigma(m)$, s^2 を次の様に定める．

$$g(m) = \int xf(x, m)\,dx\,, \tag{1.11}$$

$$\sigma^2(m) = \int (x-g(m))^2 f(x, m)\,dx\,, \tag{1.12}$$

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2. \tag{1.13}$$

例 1. (1.11) が存在するときには，$\bar{x}$ は $g(m)$ の consistent estimate であり，unbiassed estimate である．

例 2. (1.12) が存在するとき s^2 は $\sigma^2(m)$ の consistent estimate である事は明白であるが，unbiassed estimate ではない．何者：

$$\begin{aligned} E(s^2) &= E\left(\frac{1}{n}\sum_{i=1}^{n}x_i^2-\frac{1}{n^2}\left(\sum_{i=1}^{n}x_i\right)^2\right) = \frac{1}{n}\sum_{i=1}^{n}E(x_i^2)-\frac{1}{n^2}\sum_{i=1}^{n}E(x_i^2)-\frac{1}{n^2}\sum_{i\neq j}E(x_ix_j) \\ &= \frac{n-1}{n}E(x^2)-\frac{n-1}{n}E^2(x) = \frac{n-1}{n}(E(x^2)-E^2(x)) = \frac{n-1}{n}\sigma^2(m). \end{aligned} \tag{1.14}$$

故に $\sigma^2(m)$ の unbiassed estimate は $ns^2/(n-1)$ である．

例 3. $f(x, m)$ は $|x-m|$ の函數であつて $f(m, m)\neq 0$ とする．m がこの分布の中央値 (242頁, 定義 3.1) であるが，$f(m, m)\neq 0$ なる故他に中央値はない．このとき sample $(x_1, x_2, \cdots, x_{2s+1})$ の中央値 μ は m の consistent estimate である．何者 μ の密度分布を p_{2s+1} と記せば

$$\begin{cases} p_{2s+1}(x) = \dfrac{(2s+1)!}{s!\,s!}\dfrac{1}{2^{2s}}\left(1-4\left(\displaystyle\int_m^x f(x, m)\,dx\right)^2\right)^s f(x, m). \\ \lim\limits_{s\to\infty} p_{2s+1}(x) = \lim\limits_{s\to\infty}\dfrac{2s+1}{\sqrt{s}\sqrt{\pi}}\left(1-4\left(\displaystyle\int_m^x f(x, m)\,dx\right)^2\right)^s f(x, m). \end{cases} \tag{1.15}$$

從つて $\lim\limits_{s\to\infty} p_{2s+1}(m)=\infty$，$x\neq m$ ならば $\lim\limits_{s\to\infty} p_{2s+1}(x)=0$.

§ 2. Method of Maximum Likelihood

次に問題となるのは確率密度分布 $f(x; m)$ が與へられたとき，consistent statistic は存在するか，もし存在すれば如何にしてそれを見出すかといふことである．この問題に解決を與へるものは次に述べる method of maximum likelihood である．

定義 2. p ケのパラメーター $m_1, m_2, \cdots, m_p$ を含む確率密度分布 $f(x; m_1, \cdots, m_p)$ が與へられたとし，それからえられた sample を $(x_1, x_2, \cdots, x_n)$ とする．$(n \geqq p)$ この sample の生ずる確率は $\prod_{i=1}^{n} f(x_i; m_1, \cdots, m_p)$ であるが，今それを最大ならしめる $m_i = L_n^{(i)}$, $1 \leqq i \leqq p$ をとつて $(L_n^{(1)}, \cdots, L_n^{(p)})$ を $(m_1, \cdots, m_p)$ の maximum likelihood estimate といふ．パラメーターの値をその maximum likelihood estimate によつて推定する方法を method of maximum likelihood といふ．

例 4. 正規型分布

$$f(x; m, \sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \tag{N}$$

に於いて (m, σ^2) の maximum likelihood estimate を $(L^{(1)}, L^{(2)})$ とすれば

$$L^{(1)} = \bar{x}, \quad L^{(2)} = s^2 \tag{2,01}$$

である．何者

$$\Phi \equiv \log \prod_{i=1}^{n} f(x_i, m, \sigma) = \sum_{i=1}^{n} \left(-\frac{(x_i - m)^2}{2\sigma^2} - \log(\sqrt{2\pi}\,\sigma) \right) \tag{2.02}$$

とおき

$$\frac{\partial \Phi}{\partial m} = 0, \quad \frac{\partial \Phi}{\partial \sigma} = 0 \tag{2.03}$$

なる m, σ^2 を未知數とする聯立方程式を解けば，m, σ^2 の解が，夫々 $L^{(1)}, L^{(2)}$ に外ならぬわけであつて，容易に (2.01) が得られる．例 2 で述べた如く，σ^2 の unbiassed estimate は $ns^2/(n-1)$ であつた．この様に maximum likelihood estimate は，必ずしも unbiassed estimate とは一致しない．

定理 1.

(1°) $f(x, m)$ は確率 0 の x 集合を除いて m に關して微分可能.

(2°) $\displaystyle\int \frac{\partial}{\partial m} f(x, m)\, dx = 0.$

(3°) $\displaystyle u(x, m) \equiv \frac{1}{f} \frac{\partial f}{\partial m}(x, m)$ は，x に一様，m にについ て連續.

(4°) m を未知數とする方程式

$$\sum_{i=1}^{n} u(x_i, m) = 0 \tag{2.04}$$

の根 $m = M(x_1, x_2, \cdots, x_n)$ は n が何であつても常に夫々唯一つ存在する.

この四條件の下に於いて $M_n(x_1, x_2, \cdots, x_n)$ は m の consistent estimate となる.

證明 $n \to \infty$ のとき

$$\sum_{i=1}^{n} u(x_i, m)/n \tag{2.05}$$

は

$$\int u(x, m) f(x, m)\, dx = 0 \tag{2.06}$$

に法則收斂する．しかるに (3°), (4°) に依り M_n が m の近傍にあることと (2.05) が 0 の近傍にあ

ることとが同値となるから，任意の正数 ε に對して

$$\lim_{n=\infty} P(|M_n-m|>\varepsilon)=0 \tag{2.07}$$

となる．(*Q.E.D.*)

パラメーターの値を推定するには consistent statistic をとる必要があるから，これからは consistent statistic のみを考へることにするが，同じパラメーターの値を推定するにしてもその consistent estimate は色々にとり得るので，その中で最も効果的なものをえらびたいのである．

例 5. 例 3 に於ける $f(x,m)$ の二次の moment の存在を假定する．平均値が m になる事は次式より分る．

$$\int (x-m)f(x,m)\,dx=\int (x-m)f(|x-m|)\,dx=0\,.$$

標準偏差を σ と書く．次に sample $(x_1, x_2, \cdots, x_n)$ の平均値 $\bar{X}$ を m_n，中央値を μ_n とすれば，此等は何れも m の consistent estimate である．今確率變數 x の標準偏差を $\sigma(x)$ で表せば，$\sigma^2(m_n)=\sigma^2(x)/n$，$\sigma^2(\mu_n)=1/4nf^2(m,m)$.（前者は明白，後者は例 3 より計算すればよい）確率變數 $\sqrt{n}(m_n-m)$ は $n\to\infty$ のとき平均値 0 の Gauss 分布に法則收斂する (256頁，定理 1.2) が，$\sqrt{n}(\mu_n-m)$ も亦同じく $n\to\infty$ のとき平均値 0 の Gauss 分布に法則收斂する.（例 3 に於ける $p_n(x)$ より計算すればよい）

今 $f(x;m)$ として特に平均値 m，標準偏差 σ なる Gauss 分布 (N) をとる．このとき $\sigma^2(m_n)$ と $\sigma^2(\mu_n)$ とを比較すれば，$\sigma^2(m_n)/\sigma^2(\mu_n)=2/\pi$ なる故，size 100 の sample より平均値を使用することと，size $100\times\dfrac{\pi}{2}\fallingdotseq 157$ の sample より中央値を使用することとが同程度の精確さをもつことになり，sample が與へられたとき，中央値よりも平均値を使用した方が効果的であることが分る．

次に $f(x;m)$ として平均値が m の Laplace 分布をとる：$f(x;m)=e^{-|x-m|}/2$ この場合は $\sigma^2(m_n)/\sigma^2(\mu_n)=2$ であるから，size 100 の sample から平均値を使用することと，size 50 の sample から中央値を使用することとが同程度の精確さをもつことになり，Gauss 分布の場合と反對に平均値よりも中央値をとるべきである．

この例は statistic T_n について $\sqrt{n}(T_n-m)$ を作り，それが Gauss 分布へ法則收斂するとき，その Gauss 分布の標準偏差の平方によつて効果の程度を定めたのであつた．$f(x;m)$ が與へられたとき，上の意味で最も効果的なものをえらぶには如何にすればよいかといふ疑問が起るが，それも亦後で述べる様に method of maximum likelihood によつて解決される．

補助定理 確率變數列 $\{X_n\}, \{Y_n\}$ が夫々常數 $a\neq 0$，確率變數 Y に法則收斂するならば，Y_n/X_n は Y/a に法則收斂する．

證明 次の二式から容易に分る：

$$\begin{aligned} P(Y_n\leqq tX_n) &= P(Y_n\leqq tX_n,\ |X_n-a|\leqq\varepsilon)+P(Y_n\leqq tX_n,\ |X_n-a|>\varepsilon) \\ &\leqq P(Y_n\leqq t(a+\varepsilon))+P(|X_n-a|>\varepsilon)\,, \end{aligned}$$

$$P(Y_n \leqq ta) = P(Y_n \leqq ta, |X_n - a| \leqq \varepsilon) + P(Y_n \leqq ta, |X_n - a| > \varepsilon)$$
$$\leqq P(Y_n \leqq t(X_n + \varepsilon)) + P(|X_n - a| > \varepsilon).$$

定理 2.

(1°) $f(x, m)$ は確率 0 の x 集合を除いて m について二回連續的微分可能.

(2°) $\int \frac{\partial}{\partial m} f(x, m)dx = \int \frac{\partial^2}{\partial m^2} f(x, m)dx = 0.$

(3°) $u(x, m) = \frac{1}{f(x,m)} \frac{\partial f(x,m)}{\partial m}$ 及び $\frac{\partial u(x,m)}{\partial m}$ は，x に一様に m について連續.

(4°) $E(u^2) = \int u^2(x, m) f(x, m)dx < +\infty.$

この四條件の下に $\sqrt{n}(L_n - m)$ は平均値 0, 標準偏差 $1/\sqrt{E(u^2)}$ の Gauss 分布へ法則收斂する. 但し L_n は m の maximum likelihood estimate とする.

證明:

(2.08) $$g(t) = \sum_{i=1}^{n} u(x_i, t)/n$$

とおけば,

(2.09) $$g'(t) = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{f} \frac{\partial^2 f}{\partial m^2} - \left(\frac{1}{f} \frac{\partial f}{\partial m} \right)^2 \right) u(x_i, t)$$

なる故 (2.09) は，$n \to \infty$ のときには 次のものに槪收斂する：

(2.10) $$-E(u^2(x, m)) + \eta(t).$$

但し玆に

(2.11) $$\lim_{t \to m} \eta(t) = 0.$$

即ち $g(t)$ の根 $t = L_n$ は確率 0 を除けば唯一ケとなり，之は likelihood

(2.12) $$\prod_{i=1}^{n} f(x_i, t)$$

を極大にするが，更に定理 1 と同じ様にして $P(|L_n - m| > \varepsilon)$ は，$n \to \infty$ のとき，0 に收斂する.

次に

(2.13) $$0 = \sum_{i=1}^{n} u(x_i, L_n) = \sum_{i=1}^{n} u(x_i, m + L_n - m)$$
$$= \sum_{i=1}^{n} u(x_i, m) + (L_n - m) \sum_{i=1}^{n} \frac{\partial u(x_i, m + \alpha(L_n - m))}{\partial m} \quad (0 < \alpha < 1)$$

故に

(2.14) $$\sqrt{n}(L_n - m) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} u(x_i, m) \Big/ -\frac{1}{n} \sum_{i=1}^{n} \frac{\partial u(x_i, m + \alpha(L_n - m))}{\partial m}$$

が出るが，この等式の右邊の分母は，$n \to \infty$ のときには，

(2.15) $$-E\left(\frac{\partial u}{\partial m}\right) = E(u^2)$$

へ，分子は平均値 0, 標準偏差 $\sqrt{E(u^2)}$ なる Gauss 分布へ夫々法則收斂する. こゝで補助定理を使用すれば定理は證明された事になる. (*Q.E.D.*)

パラメーターの數が2個以上の場合にも定理2と大體同一の證明法によつて次の定理が證明される.

定理 2′.

(1°) $\frac{\partial^2}{\partial m_i \partial m_j} f(x; m_1, m_2, \cdots, m_p)$ は確率0の x 集合を除いては存在して連續.

(2°) $\int \frac{\partial}{\partial m_i} f(x; m_1, m_2, \cdots, m_p) dx = \int \frac{\partial^2}{\partial m_i \partial m_j} f(x; m_1, \cdots, m_p) dx = 0.$

(3°) $u_i = \frac{1}{f}\frac{\partial f}{\partial m_i}$ 及び $\frac{\partial}{\partial m_j} u$ は, x に關して一様に, $(m_1, m_2, \cdots, m_p)$ について連續.

(4°) $E(u_i^2) < +\infty$.

ここで $(L_n^{(1)}, L_n^{(2)}, \cdots, L_n^{(p)})$ を $(m_1, m_2, \cdots, m_p)$ の maximum likelihood estimate とすれば上記四條件の下に p 次元確率變數

$$\sqrt{n}\,(L_n^{(i)} - m_i) \qquad (1 \leqq i \leqq p) \tag{2.16}$$

は, $n \to \infty$ のとき, p 次元 Gauss 分布

$$\frac{\sqrt{|A|}}{(\sqrt{2\pi})^p} e^{-\frac{1}{2} y A y'} \tag{2.17}$$

に法則收斂する. 但し A は $E(u_i u_j)$ を (i, j) 元素とする p 行, p 列の行列, y は p 次元ベクター $(y_1, y_2, \cdots, y_p)$ である.

次に maximum likelihood estimate を sample の函數として表はすことを考へよう. 例へば Cauchy 分布に於て

$$\sum_{i=1}^{n} \frac{(x_i - m)}{1 + (x_i - m)^2} = 0 \tag{2.18}$$

より m を $(x_1, \cdots, x_n)$ の陽函數として表はすことは簡單ではない. しかし consistent estimate が既に見出されてゐる場合には, それから出發して次に述べる様に maximum likelihood estimate の式を近似的に解くことが出來る.

定義 3. 確率變數列 $\{x_n\}$ が0に法則收斂するとき $\underline{\lim}\, P(A \leqq n^{\alpha} X_n < B) > 0$, $AB > 0$ なる常數 A, B が存在するならば $\{x_n\}$ は order α であるといふ.

0に法則收斂する確率變數列 $\{x_n\}$ が與へられたとき, 常に order が定義されるとは限らない. しかし $\{x_n\}$ が order α であつて且つ $\beta (\neq \alpha)$ であることはない. それは次の様に證明される. $\beta > \alpha$ とし, 確率變數 $n^{\alpha} X_n$ の分布函數を $F_n(x)$ としよう: $F_n(x) = P(n^{\alpha} X_n < x)$. $\{F_n(x)\}$ より適當に部分列をえらび, 非減少函數 $G(x)$——これは必ずしも分布函數とは限らぬ——へ一様收斂させることが出來る. それを改めて $\{F_n(x)\}$ とかく. $(0, \xi)$ で $G(x)$ が連續とすれば, $F_n(x)$ は $(0, \xi)$ で $G(x)$ に一様收斂するから n を適當に大きくとれば $(F_n(D/n^{\beta-\alpha}) - F_n(C/n^{\beta-\alpha})) - (G(D/n^{\beta-\alpha}) - G(C/n^{\beta-\alpha}))$ を任意に小さくなし得る. 從つて $\underline{\lim}\, P(C \leqq n^{\beta} X_n < D) > 0$, $CD > 0$ なる C, D は存在しない. もし $G(x)$ の不連續點が0を集積點としてゐるならば, 不連續點が可附番ヶしかないことから $G(x)$ の連續點 a_n, b_n をとつて, $a_n \to 0$, $b_n \to 0$, $b_n > D/n^{\beta-\alpha} > C/n^{\beta-\alpha} > a_n$ とすることが出來る. $F_n(b_n) - F_n(a_n) \geqq F_n(D/n^{\beta-\alpha}) - F_n(C/n^{\beta-\alpha})$ であるからこの場合にも $\underline{\lim}\, P(C \leqq n^{\beta} X_n < D) > 0$,

$CD<0$ なる C, D は存在しない．(Q.E.D.)

この補助定理を利用する．今 consistent estimate を T_n, maximum likelihood estimate を L_n とすれば

(2.19) $$\sum_{i=1}^{n} u(x_i,\, T_n+L_n-T_n) = 0$$

より

(2,20) $$L_n-T_n = \frac{1}{n}\sum_{i=1}^{n} u(x_i,\, T_n) \Big/ -\frac{1}{n}\sum_{i=1}^{n} u_m(x_i,\, T_n+\alpha(L_n-T_n)). \quad (0<\alpha<1)$$

然るに L_n-T_n は 0 に収斂するから右邊の分母は $\sigma_l^2=1/-E(u_m)$ に確率收斂し，½ 以上の order を無視すれば

(2.21) $$L_n = T_n + \frac{\sigma_l^2}{n}\sum_{i=1}^{n} u(x_i,\, T_n)\,.$$

例として Cauchy の分布をとれば sample の中央値 μ_n は m の consistent estimate である(例 3) から，$\sigma_l^2=1/-E(u_m)=2$ より $\mu_n+2\sum_{i=1}^{n} u(x_i, \mu_n)/n$ を作ればこれは order ½ 以上の無視の下に maximum likelihood estimate と一致する．

§ 3. Intrinsic Accuracy, Efficiency.

定理 2 によつて，m の maximum likelihood estimate は或條件の下に標準偏差 $1/\sqrt{-E(u_m)}$ なる Gauss 分布に法則收斂することが分つた．之を利用して

定義 4. 確率密度分布 $f(x, m)$ が與へられたとき，

(3.01) $$-E\Big(\frac{\partial^2}{\partial m^2}\log f(x, m)\Big)$$

を $f(x, m)$ の intrinsic accuracy と云ひ $I(f)$ とかく．

注意 intrinsic accuracy は次の二つの著しい性質をもつ：

(3.02)
$$\text{(i)}\quad -E\Big(\frac{\partial^2}{\partial m^2}\log f(x, m)\Big) = E\Big(\Big(\frac{\partial}{\partial m}\log f(x, m)\Big)^2\Big).$$
$$\text{(ii)}\quad f(x, y, m) = f_1(x, m)\, f_2(y, m) \quad \text{ならば} \quad I(f) = I(f_1) + I(f_2).$$

何者――(i) に付いて：

(3.03) $$\text{左邊} = -\int\Big\{\frac{f_{mm}}{f} - \Big(\frac{f_m}{f}\Big)^2\Big\} f\,dx = -\int f_{mm}\,dx + \int\Big(\frac{f_m}{f}\Big)^2 f\,dx = \text{右邊}\,.$$

(ii) に付いては：

(3.04)
$$\begin{aligned} I(f) &= -E\Big(\frac{\partial^2}{\partial m^2}\log f_1 + \frac{\partial^2}{\partial m^2}\log f_2\Big) \\ &= -\iint\Big(\frac{\partial^2}{\partial m^2}\log f_1 + \frac{\partial^2}{\partial m^2}\log f_2\Big) f_1(x)\, f_2(y)\,dx\,dy \\ &= -\int f_2(y)\,dy\int\Big(\frac{\partial^2}{\partial m^2}\log f_1\Big) f_1\,dx - \int f_1(x)\,dx\,\frac{\partial^2}{\partial m^2}(\log f_2)\, f_2\,dy \\ &= I(f_1) + I(f_2)\,. \end{aligned}$$

定理 3. 定理1の四條件を滿足する確率密度分布 $f(x, m)$ から sample $(x_1, x_2, \cdots, x_n)$ を得たとする．T_1を m の estimate とし，T_1が m について微分可能なる確率密度分布 $\varphi(T_1; m)$ をもつとすれば $I(\varphi) \leqq nI(f)$.

證明 適當に $(n-1)$ ケの一價可測函數 $T_2(x_1, \cdots, x_n)$, $T_3(x_1, \cdots, x_n)$, $\cdots$, $T_n(x_1, \cdots, x_n)$ をとり，確率0を除いて，$(x_1, \cdots, x_n)$ と $(T_1, \cdots, T_n)$ とが一對一に對應する樣にする．sample $(x_1, \cdots, x_n)$ の生ずる確率は $\prod_{i=1}^{n} f(x_i, m)$ なる故 $\prod_{i=1}^{n} f(x_i, m)\, dx_i = \varphi(T_1, m) dT_1 \ \psi(T_1, T_2, \cdots, T_n, m)\, dT_2 \cdots dT_n$ とかき得る．この兩邊の intrinsic accuracy をとれば，上記注意により

$$(3.05)\qquad nE\left(\left(\frac{\partial}{\partial m}\log f\right)^2\right) = E\left(\left(\frac{\partial}{\partial m}\log \varphi\right)^2\right) = E\left(\left(\frac{\partial}{\partial m}\log \psi\right)^2\right)$$

となり，從つて $I(\varphi) \leqq nI(f)$ が證明された．

定義 5. T_1, φ を定理3で述べた通りとするとき，$I(\varphi)/nI(f)$ を T_1の efficiency といふ．

定理3及び定義5は，パラメーターの數が二ケ以上の場合にも成立する．

定理 3′. p ケのパラメーターを含む確率密度分布 $f(x; m_1, m_2, \cdots, m_p)$ が定理2′に於ける四條件を滿足するとし，$(m_1, m_2, \cdots, m_p)$ の estimate $(T_1, T_2, \cdots, T_p)$ が $m_i\ (1 \leqq i \leqq p)$ について微分可能な確率密度分布 $\varphi(T_1, T_2, \cdots, T_p, m_1, m_2, \cdots, m_p)$ をもつとする．A, B は次の如き $(i, j = 1, 2, \cdots, p)$ なる p 次行列とする．

$$(3.06)\qquad A = \left\| \frac{1}{f^2}\frac{\partial f}{\partial m_i}\frac{\partial f}{\partial m_j} \right\|,\quad B = \left\| \frac{1}{n}\frac{1}{\varphi^2}\frac{\partial \varphi}{\partial m_i}\frac{\partial \varphi}{\partial m_j} \right\|.$$

y を p 次元ヴェクター

$$(3.07)\qquad y = (y_1, y_2, \cdots, y_p)$$

とすれば，次の關係が成立つ：

$$(3.08)\qquad yAy' \geqq yBy'.$$

證明 $(n-p)$ ケの一價可測函數 $T_{p+1}(x_1, \cdots, x_n)$, $\cdots$, $T_n(x_1, \cdots, x_n)$ を適當にとつて，確率0を除いて $(x_1, x_2, \cdots, x_n)$ と $(T_1, T_2, \cdots, T_n)$ とが一對一の對應をなすとする．簡單のため $m = (m_1, m_2, \cdots, m_p)$ とかく．

$$(3.09)\qquad \prod_{i=1}^{n} f(x_i; m) = \varphi(T_1, T_2, \cdots, T_p; m)\, \psi(T_1, T_2, \cdots, T_n; m)$$

により

$$(3.10)\qquad \sum_{i=1}^{n}\frac{\partial}{\partial m_k}\log f(x_i, m)\, y_k = \left(\frac{\partial}{\partial m_k}\log\varphi + \frac{\partial}{\partial m_k}\log\psi\right) y_k \quad (1 \leqq k \leqq p)$$

そこで今

$$(3.11)\qquad \frac{\partial}{\partial m_k}\log f(x_i, m) = u_k(x_i),\quad \frac{\partial}{\partial m_k}\log\varphi = g_k,\quad \frac{\partial}{\partial m_k}\log\psi = h_k$$

とかけば，次の等式を得る：

$$(3.12)\qquad E\left\{\left(\sum_{k=1}^{p}\left(\sum_{i=1}^{n} u_k(x_i)\right) y_k\right)^2\right\} = E\left\{\left(\sum_{k=1}^{p}(g_k + h_k)\, y_k\right)^2\right\}.$$

この等式の左邊は次の如く變形される．

$$(3.13)\qquad E\left\{\sum_{k=1}^{p}\left(\sum_{i=1}^{n}u_k(x_i)\right)^2 y_k^2+\sum_{j\neq k}\left(\sum_{i=1}^{n}u_j(x_i)\right)\left(\sum_{i=1}^{n}u_k(x_i)\right)y_j y_k\right\}$$
$$=\sum_{k=1}^{p}E\left\{\left(\sum_{i=1}^{n}u_k(x_i)\right)^2\right\}y_k^2+\sum_{j\neq k}E\left\{\left(\sum_{i=1}^{n}u_j(x_i)\right)\left(\sum_{l=1}^{n}u_k(x_l)\right)\right\}y_i y_k .$$

然るに $i\neq l$ ならば j,k の如何に關せず

$$(3.14)\qquad E\{u_j(x_i)\,u_k(x_l)\}=0$$

なる故，上式は

$$(3.15)\qquad n\left[\sum_{k=1}^{p}E\left(u_k^2(x)\right)y_k^2+\sum_{j\neq k}E\left(u_j(x)\,u_k(x)\right)y_j y_k\right]=n\,E\left\{\left(\sum_{k=1}^{p}u_k(x)y_k\right)^2\right\}.$$

次に右邊をかきかへれば

$$(3.16)\qquad E\left\{\left(\sum_{k=1}^{p}g_k y_k\right)^2+\left(\sum_{k=1}^{p}h_k y_k\right)^2+\left(\sum_{k=1}^{p}g_k y_k\right)\left(\sum_{l=1}^{p}h_l y_l\right)\right\}$$
$$=E\left\{\left(\sum_{k=1}^{p}g_k y_k\right)^2\right\}+E\left\{\left(\sum_{k=1}^{p}h_k y_k\right)^2\right\}.$$

從つて

$$(3.17)\qquad \sum_{i,j}E\{g_i g_j\}\,y_i y_j\leqq n\sum_{i,j}E\{u_i u_j\}\,y_i y_j .\quad (Q.E.D.)$$

さて，

$$(3.18)\qquad \int_{-\infty}^{\infty}e^{-\frac{1}{2}yAy'}\,dy=\frac{(\sqrt{2\pi})^p}{\sqrt{|A|}}$$

であるから，上の定理により，$|B|\leqq|A|$ となる．依つて次の定義を導入する．

定義 5′. $(T_1,T_2,\cdots,T_p)$ 及び A,B を定理 3′の如く假定するとき，$|B|/|A|$ を $(T_1,T_2,\cdots,T_p)$ の efficiency といふ．

例 6. $f(x,m)=e^{-|x-m|}/2$ に於いて sample $(x_1,x_2,\cdots,x_n)$ の中央値 μ_n を m の estimate とするとき，μ_n の efficiency を計算しよう．$n=2s+1$ とする．$I(f)=1$．例3に於ける p をとれば

$$(3.19)\qquad I(p)=\frac{(s+1)(2s+1)}{s-1}\left(1-\frac{(2s)!}{2^{2s-1}(s!)^2}\right).$$

從つて μ_n の efficiency は

$$(3.20)\qquad \frac{s+1}{s-1}\left(1-\frac{(2s)!}{2^{2s-1}(s!)^2}\right).$$

この値を第1表について例示しておく．

第 1 表

s	1	2	3	4	100	∞
efficiency	0.77	0.75	0.82	0.84	0.91	1

注意 定理 3 に於ける T_1 より $S_1=\sqrt{n}(T_1-m)$ を作る．S_1 が平均値 0，標準偏差 σ の Gauss 分布へ二次の平均收斂をすると假定する．このとき T_1 の efficiency として，$n\to\infty$ ならば $\sigma^2(L)/\sigma^2(T_1)$ をとる事が出來る．但し L は m の maximum likelihood estimate とする．それ故 $n\to\infty$

のとき maximum likelihood estimate は efficient である．この様に $n\to\infty$ のときの efficiency は簡単に計算される．例 5 に於て Gauss 分布の場合ならば，sample の平均値が m の maximum likelihood estimate であるから efficiency 1. 中央値の efficiency は $\frac{2}{\pi}$. Laplace 分布ならば m の maximum likelihood estimate は中央値（$\sum_{i=1}^{n}|x_i-m|$ を最小ならしめる m は (x_i) の中央値）であるからその efficiency は 1. 平均値の efficiency は $\frac{1}{2}$ である．

例 7. χ^2 函数；観測値と理論値との差の大きさを測るものに χ^2 函数がある．今 n ケの class の各々にはいる観測値の数を $a_1, a_2, \cdots, a_n$ とし，理論値の数を $b_1, b_2, \cdots, b_n$ として $\chi^2=\sum_{i=1}^{n}(a_i-b_i)^2/b_i$ とおく

$$P(\chi^2\leqq t)=\frac{1}{2\Gamma\left(\frac{n-1}{2}\right)}\int_0^t e^{-\frac{x^2}{2}}\left(\frac{x}{2}\right)^{\frac{n-3}{2}}dx \tag{3.21}$$

を計算し，その大きさによつて適合の程度を測るのである．

今 b_i がパラメーター m を含む場合，$b_i=b_i(m)$ なるとき $\chi^2(m)$ を最小にする様に m を定め，それを T_χ とすれば，T_χ は m の consistent estimate であり，$\sqrt{n}(T_\chi-m)$ は Gauss 分布に法則収斂し T_χ の efficiency は 1 であることが證明される．(證明略)

§ 4. Sufficient Estimate, Closest Estimate.

〔1〕 **Sufficient Estimate** 例 6 で見た様に n が小さいときは，maximum likelihood estimate でも efficiency は 1 より小さいことがある．そこで efficiency 1 なる statistic があるかどうか，もしあれば如何にして見出すか等の問題が生ずる．

定義 6. efficiency 1 なる statistic を sufficient といふ．

定理 3′ の證明を見れば分る様に，ψ が $(m_1, \cdots, m_p)$ を含まないこと即ち $\partial\psi/\partial m_i=0,\ 1\leqq i\leqq p$ が $(T_1, \cdots, T_p)$ が sufficient となるための必要且つ十分な條件である．この條件は又次の様にもいへる：

「$T_i(x_1, \cdots, x_n)=T_i(x_1', \cdots, x_n'),\ 1\leqq i\leqq p$ ならば $\prod_{i=1}^{n}f(x_i; m)/\prod_{i=1}^{n}f(x_i'; m)$ が m に無關係である」

定理 4. $(m_1, \cdots, m_p)$ の sufficient estimate $(T_1, \cdots, T_p)$ が存在する爲の必要且つ十分條件は，

$$f(x, m_1, \cdots, m_p)=\exp\left(\sum_{k=1}^{q}M_kX_k+M+X\right) \tag{4.01}$$

とかゝれることである．但し $M, M_1, \cdots, M_q$ は $(m_1, \cdots, m_p)$ の函數，$X, X_1, \cdots, X_q$ は x の函數であつて

$$\sum_{i=1}^{n}X_k(x_i)=V_k(T_1, \cdots, T_p). \tag{4.02}$$

證明 簡單の爲 $p=1$ とする．$p\geqq 2$ のときも同樣な證明法で出來る．十分なことは明白．必要な事は次の如く示される：

$$L=\sum_{i=1}^{n}\log f(x_i; m) \tag{4.03}$$

とおけば

$$\partial L/\partial m=\Phi(T; m). \tag{4.04}$$

T は勿論 m を含まないから m を一つ定めれば

$$S=\sum_{i=1}^{n}g(x_i)=\sum_{i=1}^{n}f_m(x_i)/f(x_i) \tag{4.05}$$

として $T=\psi(S)$ とかける．故に

$$\frac{\partial L}{\partial m}=H(S,m),\quad \frac{\partial^2 L}{\partial x_i\partial m}=\frac{\partial H}{\partial S}\frac{\partial S}{\partial x_i}=\frac{\partial H}{\partial S}g'(x_i) \tag{4.06}$$

とかける．そこで $\partial L/\partial x_i$ は (x_i, m) の函数であるから $\partial H/\partial S$ も亦 (x_i, m) の函数となる．しかるに S は $(x_1, x_2, \cdots, x_n)$ の對稱函數であるから $\partial H(S,m)/\partial S=p(m)$. 從つて

$$H(S,m)=p(m)S+q(m)\,. \tag{4.07}$$

即ち

$$\sum_{i=1}^{n}\frac{\frac{\partial f}{\partial m}(x_i,m)}{f(x_i,m)}=p(m)\sum_{i=1}^{n}g(x_i)+q(m)\,. \tag{4.08}$$

特に $x=x_i\ (1\leqq i\leqq n)$ を考へれば

$$\log f(x,m)=M_1(m)g(x)+M(m)+h(x)\,.\quad (Q.E.D.) \tag{4.09}$$

例 8. (1°) Gauss の分布

$$f(x;m,\sigma)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}} \tag{N}$$

に於いて (m,σ) が未知のとき，

$$\begin{cases} T_1=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}x_i \quad (=\bar{x}), \\ T_2=\sqrt{\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i-T_1)^2} \quad (=s_2) \end{cases} \tag{4.10}$$

が (m,σ) の sufficient estmate である．その理由は

$$\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}x_i',\quad \sum_{i=1}^{n}x_i^2=\sum_{i=1}^{n}x_i'^2 \tag{4.11}$$

より次の等式を得るからである：

$$L_1=\sum_{i=1}^{n}\log\frac{f(x_i,m,\sigma)}{f(x_i',m,\sigma)}=+\frac{1}{2\sigma^2}\left(\sum x_i'^2-\sum x_i^2\right)+2m\left(\sum x_i-\sum x_i'\right)=0. \tag{4.12}$$

次に σ を既知，m を未知としよう．このときには

$$\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}x_i' \tag{4.13}$$

だけから

$$L_1=\sum(x_i'^2-x_i^2)/2\sigma^2 \tag{4.14}$$

となり，これは m を含まないから T_1 が m の sufficient estimate である．m が既知，σ が未知のときには T_2 が σ の sufficient estimate となることは L_1 の形から明白である．

(2°) Pearson Type III.

$$f(x;m_1,m_2,m_3)=\frac{1}{m_2\Gamma(m_3+1)}\left(\frac{x-m_1}{m^2}\right)^{m_3}e^{-\frac{(x-m_1)}{m_2}} \tag{4.15}$$

に於て，m_1 が既知ならば，m_2, m_3 の兩方或は何れか一方が未知のとき sufficient estimate が存在することは容易に證明される．もし m_1 が未知ならば sufficientes timate は存在しない．何者 m_1 の sufficient estimate があれば

$$\frac{\partial^2 f}{\partial x \partial m_1} = \sum_{k=1}^{q} \frac{\partial M_k}{\partial m_1} \frac{dX_k}{dx}. \tag{4.16}$$

$y=-m_1$ とおけば，この等式の左邊は $F(x+y)$ の形となる．從つて $\{A_j(x)\}$, $1 \leqq j \leqq m \leqq q$ を一次獨立として次の如くかける．

$$F(x+y) = \sum_{k=1}^{m} A_k(x) B_k(y). \tag{4.17}$$

これを y に關して m 回微分して出來る $(m+1)$ ケの式より $A_1(x), \cdots, A_m(x)$ を消去して $y=0$ とおけば $F(x)$ は常數係數の同次線型微分方程式の解となる．それ故 $P_k(x)$ を多項式として

$$F(x) = \sum_{k=1}^{N} P_k(x) e^{\alpha_k x} \tag{4.18}$$

となる筈であるが，今は $\partial^2(\log f)/\partial x \partial m_1 = m_3/(x-m_1)^2$ であるから $F(x)=m_3/x^2$ となつて結局 m_1 の sufficient estimate のないことが分つた．

〔2〕**Closest Estimate** 同一のパラメーターの値に關する二つの consistent estimate があつたとき何れを採用すべきかについて，以上述べた方法は R. A. Fisher によるものである．Pitman は sufficient estimate を利用して異つた立場から一つの方法を提出したが主定理は未だ證明されてゐない．

定義 7. T_1, T_2 を $g(m)$ の consistent estimate とする．凡ての m について，$P(|T_1-g(m)| \leqq |T_2-g(m)|) > \frac{1}{2}$ ならば T_1 は T_2 より closer であるといふ．T_2 を任意にとつて常に T_1 が T_2 より closer ならば，T_1 を $g(m)$ の closest estimate といふ．

注意 T_1 が T_2 より closer, T_2 が T_3 より closer, T_3 が T_1 より closer なることもある．

m の sufficient estimate T が存在するとき，T は m に關して sample の有する性質を最も多く保有せるものと考へられるから，closest estimate が T より導かれると豫期するのは當然であらう．今次の事實，「確率變數 X の中央値を μ とし，ν を $\frac{1}{2}(\mu+\nu)$ が中央値でないようにとれば $P(|X-\mu| \leqq |X-\nu|) > \frac{1}{2}$」の成立することを考へれば，closest estimate の中央値が m に等しい事が望ましい．それ故，sufficient estimate T の中央値を $\psi(m)$ とすれば，$\psi(m)$ が單調函數なるとき $\psi^{-1}(T)$ は中央値 m の sufficient estimate である．この $\psi^{-1}(T)$ が m の closest estimate であるといふのが Pitman の豫想である．

例 9. Gauss 分布 (N) (例 8) に於いて m, σ を未知とする．size n の sample $(x_1, \cdots, x_n)$ より (4.10) 及び (4.11) を作れば $(T_1, T_2/n)$ は (m, σ^2) の sufficient estimate である．(例 8, 1°) T_2/σ^2 の密度分布は自由度 $(n-1)$ の χ^2- 分布と同一である：

$$P(T_2/\sigma^2 < t) = \frac{1}{2\Gamma\left(\frac{n-1}{2}\right)} \int_0^t e^{-\frac{x}{2}} \left(\frac{x}{2}\right)^{\frac{n-3}{2}} dx. \tag{4.19}$$

この中央値を k_{n-1} で表せば σ^2 の closest estimate は T_2/k_{n-1} である．但し $k_{n-1} \fallingdotseq n-2/3+0.09/n$.

σ^2 の estimate として通常 unbiassed estimate $S_2/n-1$ をとるが，これは $P(S_2/n-1<\sigma^2)>1/2$ なる故小さ過ぎるといふことが出來る．

この例及び他の特別な場合について上記豫想の正しいことが證明されてゐる．(證明略)

§ 5. f(x; m) の range が m に關係する場合

今迄は「$f(x; m)$ の range は m に關係しない」といふ假定をおいたがこの § では上の假定を取去つて見る．この場合定理 2, 3 が必ずしも成立しないことは次の簡單な例で分る．

(5.01) $$\int_{-\infty}^{m} f(x; m)\,dx = 1\,,\quad g(m) = \frac{1}{n}\sum_{i=1}^{n}\log f(x_i; m)$$

とおけば $n\to\infty$ のとき $g'(m)\to -f(m, m)\leqq 0$. 從つて $f(m, m)>0$ と假定すれば $g'(m)<0$ であるから $m=\underset{1\leqq i\leqq n}{\mathrm{Max}}(x_i)$ のとき $g(m)$ は最大となる．$L_n=\underset{1\leqq i\leqq n}{\mathrm{Max}}(x_i)$ の分布密度は

(5.02) $$p_n(x) = n\left(\int_{-\infty}^{x} f(x, m)\,dx\right)^{n-1} f(x; m).$$

$x=m$ のとき $p_n\to\infty$, $x\neq m$ のとき $p_n\to 0$. 從つて L_n は consistent である．しかし $\sqrt{n}(L_n-m)$ は Gauss 分布に收斂するとは限らぬ．例へば $0\leqq x\leqq m$ なるとき $f(x, m)=1/m$, その他の場合 $f(x,m)=0$ とすれば $n(L_n-m)$ の分布密度は $e^{\frac{x}{m}}dx/m(-\infty<x<0)$ に收斂する．即ち order 1 である．このとき m_n を sample の平均の二倍，μ_n を sample の中央値の二倍とすれば $\sqrt{n}(m_n-m)$, $\sqrt{n}(\mu_n-m)$ は共に Gauss 分布に收斂するが $\mathrm{Max}(x_i)$ の方が遙かに効果的なことは order を考へれば明白であらう．$I(f)=1/m^2$, $I(p)=n^2m^2$. 故に $I(p)>nI(f)$.

次に又

(5.03) $$f(x, m) = e^{-(x-m)},\quad m<x<\infty$$

を考へれば

(5.04) $$E\left\{\left(\frac{\partial}{\partial m}\log f\right)^2\right\} = E(1) = 1,\quad E\left\{\frac{\partial^2}{\partial m^2}\log f\right\} = E(0) = 0.$$

それ故，定義 4, 注意 i) の等式は成立しない．

次に

(5.05) $$\int_a^b f(x, m)\,dx = 1$$

を m で微分すれば

(5.06) $$J(f) = \int_a^b \frac{\partial f}{\partial m}dx = a'(m)f(a(m), m) - b'(m)f(b(m), m)$$

となつて一般に $J(f)\neq 0$. 然るに $(x_1, x_2, \cdots, x_n)$ の likelihood L については

(5.07) $$E\left\{\left(\frac{\partial}{\partial m}\log L\right)^2\right\} = E\left\{\left(\frac{\partial}{\partial m}\log(\Pi f(x_i, m))\right)^2\right\}$$
$$= nE\left\{\left(\frac{\partial}{\partial m}\log f(x, m)\right)^2\right\} + n(n-1)J^2(f).$$

$f(x, m)$ の intrinsic accuracy $I(f)$ として

(5.08) $$I(f) = E\left\{\left(\frac{\partial}{\partial m}\log f\right)^2\right\}$$

をとる事にすれば，sample $(x_1, x_2, \cdots, x_n)$ の intrinsic accuracy は $nI(f)+n(n-1)J^2(f)$ となる．

他の例として

$$f(x, m)=\frac{2x}{2m+1}, \quad (m \leqq x \leqq m+1) \tag{5.09}$$

をとれば

$$E\left\{\left(\frac{\partial}{\partial m}\log L\right)^2\right\}=\frac{4n^2}{(2m+1)^2}. \tag{5.10}$$

$$T=\underset{1\leqq i\leqq n}{\mathrm{Max}}\{x_i\} \tag{5.11}$$

をとれば，その密度分布は

$$p(x, m)=\frac{2nx(x^2-m^2)^{n-1}}{(2m+1)^n}. \tag{5.12}$$

これより

$$E\left\{\left(\frac{\partial}{\partial m}\log p\right)^2\right\}>nI+n(n-1)J^2. \tag{5.13}$$

この様に $f(x; m)$ の range に m が關係する場合にも有効な intrinsic accuracy の定義は未だ見付からない．

II. 區間による推定法

以上に於ては常に一ケの値によつてパラメーターの値を推定することを試みた．しかし如何なる estimate を採用しても，その値が丁度パラメーターの値に等しいといふことは先づ不可能である．それでよく行はれてゐるのは estimate T の標準偏差 $\sigma(T)$ をとつて $T\pm c\sigma(T)$ (c は常數) によつてパラメーターの値を表はす方法である．かくしてパラメーターの値を推定するのに一ケの値を使用するのではなく，區間を使用する試みが問題となる．しかし之は problem of estimation といふよりむしろ所謂 problem of testing statistical hypothesis に密接な關係を有するものである故，簡單に紹介することとするが，それは主として Neyman の理論である．

sample の空間 S——その點即ち sample $(x_1, \cdots, x_n)$ を x で表はす——には p ケのパラメーター $m=(m_1, \cdots, m_p)$ を含む確率分布 $P(A; m)$, $A\subset S$ が與へられてゐるとする．更に P は密度 $p(x; m)$ を有するとする．今 sample $(x_1, \cdots, x_n)$ が得られたとし，之から一ケのパラメーター m_1 を推定する方法を述べよう．ここでは $x_1, \cdots, x_n$ が相互に獨立であることは假定してないし，又 $f(x, m)$ の range も——それが考へられる場合——に必ずしもパラメーターに無關係ではないことは注意を要する．

定義 1. $0\leqq\alpha\leqq 1$ が與へられたとし，S で定義され實數値をとる二つの一價函數 $\underline{m_1}(x)\leqq\overline{m_1}(x)$ を適當に定めて，$m_2, \cdots, m_p$ に無關係に等式 $P([x; \underline{m_1}(x)\leqq m_1\leqq\overline{m_1}(x)]; m)=\alpha$ が成立したとする．このとき $\underline{m_1}(x)$, $\overline{m_1}(x)$ を夫々 m_1 の lower, upper confidence limit, α を confidence coefficient といひ，區間 $[\underline{m_1}(x), \overline{m_1}(x)]$ を α に對する confidence interval といつて $\sigma(x)$ くとか．

定義 2. $0\leqq\alpha\leqq 1$ とする．m_1 を任意に與へたとき $A(m_1)\subset S$ が次の四條件を滿足する樣にとれるとする．

(i). 任意の x に對して $x\in A(m_1)$ なる如き m_1 が存在する.

(ii). $P(A(m_1');m)$ が凡ての (m_1',m) について定義されその値は α に等しい.

(iii). $m_1^{(1)}<m_1^{(2)}<m_1^{(3)}$ ならば $A(m_1^{(1)})\cdot A(m_1^{(3)})\subset A(m_1^{(2)})$.

(iv). $\prod_{m_1^{(1)}<m_1<m_1^{(2)}} A(m_1)\subset A(m_1^{(1)})\,A(m_1^{(2)})$.

このとき $A(m_1)$ を α に對する m_1 の region of acceptance といふ.

定理 1. $0\leqq\alpha\leqq 1$ を與へたとき, 凡ての $x\in S$ について α に對する confidence limit $(\underline{m_1}(x), \overline{m_1}(x))$ を作ることと凡ての m_1 について α に對する region of acceptance $A(m_1)$ を作ることとは同等である.

證明 $A(m_1)=[x;\underline{m_1}(x)\leqq m_1\leqq\overline{m_1}(x)]$; $\underline{m_1}(x)=\inf.[m_1;x\in A(m_1)]$, $\overline{m_1}(x)=\sup.[m_1;x\in A(m_1)]$ より容易に證明される.

例 1. sample の空間を二次元とし, 與へられた確率密度を「$p(x_1,x_2;m)=m^{-2}$, $0<x_1,x_2<m$; $p(x_1,x_2;m)=0$, その他の場合」とする. confidence intervals を定めて見よう. 前定理によつて regions of acceptance を定めればよいがそのとり方には色々ある. そこで $A_1(m)$, $A_2(m)$, $A_3(m)$ を次の様に定める. $A_1(m)=[(x_1,x_2);\beta m\leqq x_1,x_2\leqq m]$, $0<\beta<1$. $A_2(m)=[(x_1,x_2);\gamma\leqq x_1+x_2\leqq m+\gamma]$, $0<\gamma<m$. $A_3(m)=[(x_1,x_2);\Delta m\leqq \mathrm{Max}(x_1,x_2)\leqq m]$, $0<\Delta<1$. 先づ $A_1(m)$: これが region of acceptance とならない事は, 點 (x_1,x_2), $x_2<\beta x_1$ に對しては如何に m をとつても $\beta m\leqq x_1$, $x_2\leqq m$ が成立しないことから分る. 次に $A_2(m)$: これが region of acceptance になることはすぐ分る. confidence coefficient を α とすれば $(m^2-\gamma^2)/m^2=\alpha$ なる故 $\gamma=\sqrt{1-\alpha}\,m$. 從つて $\underline{m}(x)=(x_1+x_2)/(1+\sqrt{1-\alpha})$, $\overline{m}(x)=(x_1+x_2)/\sqrt{1-\alpha}$. confidence interval の長さは $(x_1+x_2)/\sqrt{1-\alpha}(1+\sqrt{1-\alpha})$. 終りに $A_3(m)$: これも region of acceptance なる事は容易に分る. $\Delta=\sqrt{1-\alpha}$, $\underline{m}(x)=\mathrm{Max}(x_1,x_2)$, $\overline{m}(x)=\mathrm{Max}(x_1,x_2)/\sqrt{1-\alpha}$, confidence interval の長さは $(1-\sqrt{1-\alpha})\mathrm{Max}(x_1,x_2)/\sqrt{1-\alpha}$. 上の二つの confidence interval の長さを比較すれば $(x_1+x_2)/\sqrt{1-\alpha}(1+\sqrt{1-\alpha})>(x_1+x_2)(1-\sqrt{1-\alpha})/\alpha\sqrt{1-\alpha}\leqq(1-\sqrt{1-\alpha})\mathrm{Max}(x_1,x_2)/\sqrt{1-\alpha}$. 故に $A_2(m)$ をとるより $A_3(m)$ をとつた方がよい.

上例ではパラメーターが一ケであつたが二ケ以上になると confidence interval の作り方は非常に難しくなる. それは $A(m_1)$ を $m_2,\cdots,m_p$ に無關係に定めなくてはならないからである. しかし次の定理が成立するときは後例で見る如く簡單に region of acceptance を作ることが出來る.

定理 2.

(i). x について $S(<n)$ ケの一價可測函數 $T_1(x),\cdots,T_S(x)$ が獨立である.

(ii). $p(x;m)=p(T_1,\cdots,T_S;m)f(x;m_1)$. 但し $p(T_1,\cdots,T_S;m)=p([x;T_1(x)=T_1,\cdots,T_S(x)=T_S];m)$.

この二條件の下に次の結果が生ずる.

(1°) $B=[x;T_1(x)=c_1,\cdots,T_S(x)=c_S]$, $C\subset B$ とするとき $P_B(C;m)=P(C;m)/P(B;m)$ は $m_2,\cdots,m_p$ に無關係である.

(2°) $0\leqq\alpha\leqq1$ が與へられたとし, $m_1=m_1^0$ を定めて $P_B(C;m_1^0,m_2,\cdots,m_p)=\alpha$ なる如く C を

とれば $P([C\times(T_1,\cdots,T_S)];m_1^0,m_2,\cdots,m_p)=\alpha$.

證明 容易に出來る．省略する．

例 2. 「$0\leqq x_1+x_2\leqq m_1$, $0\leqq x_1,x_2\leqq m_1$ ならば $p(x_1,x_2;m_1,m_2)=2/m_1^2-m_1m_2+3m_2x_1$．他の場合は $p(x_1,x_2;m_1,m_2)=0$」なる確率分布が與へられてゐるとき region of acceptance を見出さう．$0\leqq x_1\leqq m_1$ のとき $p(x_1)=\int_0^{m_1-x_1}p(x_1,x_2;m_1,m_2)dx_2=p(x_1,x_2;m_1,m_2)(m_1-x_1)$. その他の場合 $p(x_1)=0$. 從つて $0\leqq x_1,x_2\leqq m_1$, 且 $0\leqq x_1+x_2\leqq m_1$ なるとき $f(x|m_1)=(m_1-x_1)^{-1}$. その他のとき $f(x|m_1)=0$ とすれば $p(x;m)=p(x_1;m)f(x|m_1)$. それ故 m_1 の値を一つ定め定理2に於ける $C(x_1)$ を見出せばよいがそれは簡單である．實際 $0\leqq x_2\leqq m_1-x_1$ ならば $p_{x_1}(x_2;m)=p(x;m)/p(x_1;m)=(m_1-x_1)^{-1}$. その他のとき $p_{x_1}(x_2;m)=0$. 即ち x_1 を與へたならば x_2 の確率分布は一様であるから $C(x_1)=[x_2;b(m_1-x_1)\leqq x_2\leqq(b+\alpha)(m_1-x_1)]$, $b\leqq 1-\alpha$ とすればよい．それ故 $A_2(m_1)=[x;b(m_1-x_1)\leqq x_2\leqq(b+\alpha)(m_1-x_1)]$ とおけば，これが region of acceptance になることは明白である． confidence limits は，$\underline{m_1}(x)=x_1+x_2/(b+\alpha)$, $\overline{m_1}(x)=x_1+x_2/b$. confidence interval の長さは $\alpha x_2/b(b+\alpha)$. これは $b=1-\alpha$ のとき最短で從つて最も効果的である．$C(x_2)$ のとり方は勿論他にもある．例へば $0\leqq x_1<m_1/2$ ならば $C(x_1)=[(x_1,x_2);(1-\alpha)(m_1-x_1)\leqq x_2\leqq m_1-x_1]$, $x_2=m_1/2$ のとき $C(x_1)=[(x_1,x_2);0\leqq x_1\leqq x_2]$, $m_1/2<x_1\leqq m_1$ ならば $C(x_1)=[(x_1,x_2);0\leqq x_2\leqq\alpha(m_1-x_1)]$. 次に $A_3(m_1)=[(x_1,x_2);0\leqq x_1\leqq m_1,\ x_2\in C(x_1)]$ とすれば $A_3(m_1)$ は $\alpha\leqq 1/2$ のとき region of acceptance となる．$\alpha<1/2$ のときにさうでない事は $\alpha<x_1/x_2<1-\alpha$ なる (x_1,x_2) をとればすぐ分る．$\alpha>1/2$ のときの confidence limits は：$x_1\leqq x_2$ ならば $\underline{m_1}(x)=x_1+x_2$, $\overline{m_1}(x)=x_1+x_2/(1-\alpha)$; $\alpha x_1\leqq x_2\leqq x_1$ ならば $\underline{m_1}(x)=2x_1$, $\overline{m_1}(x)=x_1+x_2/(1-\alpha)$; $(1-\alpha)x_1\leqq x_2\leqq\alpha x_1$ ならば $\underline{m_1}(x)=x_1+x_2/\alpha$, $\overline{m_1}(x)=x_1+x_2/(1-\alpha)$, $x_2\leqq(1-\alpha)x_1$ ならば $\underline{m_1}(x)=x_1+x_2/\alpha$, $\overline{m_1}(x)=2x_1$. これより confidence interval の長さもすぐ計算出來るが，特別の場合として，$\alpha x_1=x_2$ のときは $(2\alpha-1)/\alpha(1-\alpha)$ となつて $A_2(m)$ より短いが $qx_1=x_2$, $0<q<\alpha(1-\alpha)/(1-\alpha+\alpha_2)$ のときは $x_1-x_2/\alpha>\alpha x_2/(1-\alpha)$ となつて $A_2(m)$ より長い．

定義 3. $\delta_1=[\delta_1(x);x\in S]$ を $0\leqq\alpha\leqq 1$ に對する system of confidence intervals とする．他の任意の system of confidence intervals δ_2 について，(m,m') の如何に拘らず $P([x;m_1\in\delta_1(x)];m_1')\leqq P([x;m_1\in\delta_2(x)];m_1')$ となるとき δ_1 を shortest system of confidence intervals といふ．

shortest system of confidence intervals の構成法は：

(1°). m_1 の一つの値 m_1^0 に對し $A(m_1^0)$ を $m_2,\cdots,m_p$ に無關係に，且 $P(A(m_1^0);m_1^0,m_2,\cdots,m_p)=\alpha$ (α は $0\leqq\alpha\leqq 1$ なる常數) となる様に定める．

(2°). $\overline{m_1}$ を 1° を滿足する凡ての $A(m_1^0)$ の中で $m_1^0\neq\overline{m_1}$ なるとき $P(A(m_1^0);\overline{m_1})$ が最小となる様に $A_0(m_1^0)$ をえらぶ．

(3°). 上記 $A_0(m_1^0)$ が $\overline{m_1}$ を動かしても尙 $P(A(m_1^0);\overline{m_1})$ を最小にして居り，且 m_1^0 を動かしても α に對する region of acceptance になつてゐるとする．

かく定められた $\{A_0(m_1^0)\}$ は system of regions of acceptance であつてそれより作つた system of confidence intervals が shortest なることは明白である．上記 (1°), (2°) は肯定的に解かれるが，(1°), (2°) を満足するものが必ずしも (3°) を満足しないのである．即ち $P(A_0(m_1^0);\overline{m_1})$ を最小にする $A_0(m_1^0)$ は $\overline{m_1}$ をかへるとき一般にはこの性質を保有しないのである．

補助定理 「$\int_A p(x,m)\,dx=$ 一定」なる條件の下に $\int_A p(x,m')\,dx$ を極小ならしめるための必要十分條件は $A=[x;p(x;m')\leqq cp(x;m)]$ とかけることである．

證明 (省略).

定理 3.

(1°). 一ケのパラメーターを含む確率密度分布 $p(x;m)$ が次の三條件を満足する．

(i). sample の空間 S に於て微分可能.

(ii). $\partial p(x;m)/\partial m$ が存在して連續，且 $\neq 0$.

(iii). $\int_S \partial p(x;m)/\partial m\,dx=0$.

(2°). m の二ケの定つた値 m_1, m_2 に關し，$P(A(m_1);m_1)=P(A;m_1)=\alpha$，$P(A(m_1);m_2)\leqq P(A;m_2)$ を満足する $A, A(m_1)\subset S$ が存在する．

(3°). $A(m_1)$ の境界上の少くとも一點に於て $p(x;m_1)\neq 0$.

(1°), (2°), (3°) なる條件の下に次の結果が成立する．m の値を適當にとり——之を m_3 とする——更に $B\subset S$ を適當に定めて $P(B;m_1)=\alpha$, $P(A(m_1);m_3)>P(B;m_3)$ とすることが出來る．

證明 $P(B;m_3)=\alpha$ なる任意の $B\subset S$ 及び任意の m_3 について常に $P(A(m_1);m_3)\leqq P(B;m_3)$ が成立すると假定しよう．補助定理により x が $A(m_1)$ の夫々內點・境界點・外點なるとき $p(x;m)\lesseqqgtr k(m)\,p(x;m_1)$ を満足する $k(m)$ の存在が分る．次に $A(m_1)$ の境界上の一點 x_0 を $p(x;m_1)>0$ となる様にとれば $k(m)=p(x_0;m)/p(x_0;m_1)$．書直せば $k(m)=1+(m-m_1)p_m(x_0,m_1+\theta(m-m_1))/p(x_0;m_1)$, $0<\theta<1$. 一方 $x_1\in A(m_1)$ に對して $p(x_1;m)\leqq k(m)\,p(x_1;m_1)$．或は $(m-m_1)(p_m(x_1;m_1+r(m-m'))-p_m(x_0;m_1+\theta(m-m_1))\,p(x_1;m_1)/p(x_0;m_1))\leqq 0$, $0<r<1$. 故に $p_m(x_1;m_1)-p(x_1;m_1)/p(x_0;m_1)=0$. (もし左邊 $\neq 0$ ならば，これは $p_m(x;m)$ の連續性により m が m_1 に十分近いとき一定符號をもつ．しかるに $m-m_1$ は正にも負にもなり得るから不合理.) 同様にしてこの等式は $x_1\bar{\in} A(m_1)$ でも成立し從つて任意の x について $p_m(x;m_1)=p_m(x_0;m_1)\,p(x;m_1)/p(x_0;m_1)$．又假定 (1°), (iii) 及び

$$\int_S p_m(x,m_1)\,dx=\frac{p_m(x_0;m_1)}{p(x_0;m_1)} \tag{1.01}$$

によつて $p_m(x_0;m_1)=0$ となつて不合理である．

この定理で分る様に $p_m(x;m)$ が十分正則な性質をもてば shortest system of confidence intervals は存在しないのである．それで最も効果的と考へられる system of confidence intervals の特徴を如何に云表はすかといふことが問題となる．

定義 4. $0\leqq\alpha\leqq 1$ に對する system of confidence intervals の集合 $\{\delta(x)\}$ に於いて，$\partial P([x;m_1\in\delta(x)],m)/\partial m_1=0$ を満足するものの中，$\partial^2 P([x;m_1\in\delta(x)],m)/\partial m_1^2$ を極小にするものを

short unbiassed system of confidence intervals といふ．

以後パラメーターの數を一ケとし，$p(x;m)$ は次の四條件を滿足するものと假定する．

(1°). $p(x,m)$ は m について微分可能．

(2°). $\frac{\partial}{\partial m}\log p$ は m を如何に定めても S で恒等的に 0 となることはない．

(3°). $\int_S\left(\frac{\partial}{\partial m}\log p\right)^2 p(x,m)\,dx<\infty$.

(4°). $\frac{\partial^2}{\partial m^2}\log p=u(m)+v(m)\frac{\partial}{\partial m}\log p$.

注意 (4°) は $x_1,\cdots,x_n$ が獨立なるとき sufficient estimate が存在するための必要且十分條件である：實際

$$p(x;m)=\prod_{i=1}^{n}f(x;m) \tag{1.03}$$

とおけば $\partial^2(\log f(x;m))/\partial m^2=u_0(m)+v_0(m)\,\partial(\log f(x;m))/\partial m$. $\partial(\log f)/\partial m=g$ とおけば $g_m=u_0+v_0 g$. $\therefore\ (g_x)_m=v_0 g_x\rightarrow\partial(\log g_x)/\partial m=v_0\rightarrow g_x=MX\rightarrow g=M_1X_1+N_1\rightarrow\log f=M_1X_1+M_2+X_2\rightarrow f=\exp(M_1X_1+M_2+X_2)\rightarrow\partial^2\log f/\partial m^2=u(m)+v(m)\,\partial(\log f)/\partial m$.

次に short unbiassed system of confidence intervals の條件は，上の $p(x;m)$ を使用し次の如くおく．

$$\varphi=\left(\frac{\partial}{\partial m}\log p(x;m)\right)_{m=m_1},\quad \varphi'=\left(\frac{\partial^2}{\partial m^2}\log p(x;m)\right)_{m=m_1} \tag{1.04}$$

そのとき

(i). $\int_{A(m_1)}p(x,m_1)\,dx=\alpha$,

(ii). $\int_{A(m_1)}\varphi p(x;m_1)\,dx=0$,

(iii). $\int_{A(m_1)}(\varphi'+\varphi^2)p(x;m_1)\,dx=$ 極小 と表はされる．

故に補助定理により $A(m_1)=[x;\ \varphi'+\varphi^2\leqq a\varphi+b]$ となる．ここで a,b は (i),(ii) より定まる常數である．任意の m_1 について上の條件を滿足する $A(m_1)$ が作られたとし，更に $\sum_{m_1}A(m_1)=S$ が成立したとすれば，これは region of acceptance であり，從つて short unbiassed system of confidence intervals が構成されたことになる．

定理 4. 任意の m_1 に對して適當に常數 a,b をとれば $A(m_1)=[x;\ \varphi'+\varphi^2\leqq a\varphi+b]$ としたとき

$$\int_{A(m_1)}p(x;m_1)\,dx=\alpha\,,\quad \int_{A(m_1)}\varphi p(x;m_1)\,dx=0 \tag{1.05}$$

證明 $p(x;m)=\exp(M_1(m)+M_2(m)\,t(x)+S(x))$ とおけば $\varphi=M_1'(m_1)+M_2'(m_1)\,t(x)$. 次に

$$\int_S p(x;m)\,dx=1 \tag{1.06}$$

の兩邊を m で微分する．

$$\frac{d}{dm}\int_S p(x;m)\,dx=M_1'(m)+M_2'(m)\int_S t(x)\,p(x;m)\,dx=0. \tag{1.07}$$

もし $M_2'(m)=0$ ならば $M_1'(m)=0$. 即ち $\varphi(x;m)=0$. これは假定に反する．故に $M_2'(m)\neq 0$. 從つて $\varphi'+\varphi^2\leqq a\varphi+b$ は $t^2\leqq a_1t+b_1$ 又は $t_1\leqq t\leqq t_2$ と同値になる．結局 $p(x;m)=p(t;m)$ として

(1.08) $$\int_{t_1}^{t_2} p_1(t, m_1)\, dt = \alpha, \quad \frac{1}{\alpha}\int_{t_1}^{t_2} t p_1(t, m_1)\, dt = -\frac{M_1'(m_1)}{M_2'(m_1)}$$

なる様に t_1, t_2 を定めればよいのである.

定理 5. $A(m_1)$ を定理 4 の如くにとる.

(1.09) $$\int_{B(m_1)} p(x; m_1)\, dx = \alpha, \quad \int_{B(m_1)} \varphi p(x; m_1)\, dx = 0$$

を滿足する領域 $B(m_1)$ を如何にとつても任意の m に對して,

(1.10) $$\int_{A(m_1)} p(x; m)\, dx \leqq \int_{B(m_1)} p(x; m)\, dx.$$

證明

(1.11) $$\int_{B(m_1)} p(x; m)\, dx = \alpha, \quad \int_{B(m_1)} \varphi p(x; m)\, dx = 0$$

なる條件の下に

(1.12) $$\int_{B(m_1)} p(x; m)\, dx$$

を最小ならしめる $B(m_1)$ 即ち $A(m_1)$ を見出すことが問題であるが, 上の二條件を滿足する樣に常數 c, d を定めて, $A(m_1) = [x; p(x, m) \leqq (c\varphi + d) p(x; m_1)]$ とかければよいことは補助定理から分る. $P = M_1(m) - M_1(m_1)$, $Q = M_2(m) - M_2(m_1)$ として $p(x; m) \leqq (c\varphi + d) p(x, m_1)$ をかきかへれば $e^{P+Qt(x)} \leqq c\varphi + d = c_1(m_1)\, t(x) + d_1(m_1)$. c, d を定めることと c_1, d_1 を定めることは同値なる故, (t_1, t_2) を定理 4 の如く定めて $e^{P+Qt_1} = c_1 t_1 + d_1$, $e^{P+Qt_2} = c_1 t_2 + d_1$ より c_1, d_1 を決定すればよい. しかるに $y(t) = e^{P+Qt} - c_1 t - d_1$ とおけば $y''(t) = Q^2 e^{P+Qt} > 0$. 從つて $y(t)$ は $t_1 < t < t_2$ のとき負となる. (*Q.E.D.*)

この定理は簡單にいへば short unbiassed system of confidence intervals が適當な條件の下では或る意味で shortest なることを主張するものである. 適當な條件とは $\sum_{m_1} A(m_1) = S$ が成立することであり, 或る意味で shortest とは

(1.12) $$\int_{B(m_1)} p_m(x; m_1)\, dx = 0$$

を滿足する領域 $B(m_1)$ の中で shortest といふことである. $\sum_{m} A(m) = S$ の一つの十分條件を與へるものとして

定理 6. 任意に x を與へたとき $M_1'(m_1) + M_2'(m_1)\, t(x) = 0$ を滿足する $m_1(x)$ が存在するならば $\sum_{m} A(m) = S$.

證明 $x \in S$ に對して上の樣に m_1 を定め, $A(m_1)$ を作れば $x \in A(m_1)$ となる. それは $t_1 \leqq t(x) \leqq t_2$ より明白である.

注意 上定理で定めた $m_1(x)$ が maximum likelihood estimate である. 何者:

(1.13) $$\frac{\partial}{\partial m} p(x; m) = (M_1' + M_2' t)\, p, \quad \frac{\partial^2}{\partial m^2} p = (M_1'' + M_2'' t + (M_1' + M_2' t)^2)\, p.$$

故に上定理の m_1 を代入すれば

$$(1.14)\qquad \left[\frac{\partial}{\partial m}p(x, m)\right]_{m=m_1} = 0\,, \qquad \left[\frac{\partial^2}{\partial m^2}p(x; m)\right]_{m=m_1} = \frac{M_2''M_2'-M_1'M_2''}{M_2'} < 0\,.$$

(この關係は $\int_S (M_1'+M_2't)^2 p\,dx = -\dfrac{M_1''M_2-M_1'M_2''}{M_2'}$ より出る)

注意 a を如何に與へても $p(x; m)$ を最大にする樣な m が常に $\delta(x)$ に含まれることが分る.

脚　　註

* 北川敏男：獨立確率變數の理論，日本數學物理學會誌，第14卷，240頁，定義 2.1 を參照．以下も同じ形式で上記論文を參照する．