⑰ I. Amount of Information について

渡辺壽夫

I. まへがき

電気通信工学に関する統計的研究に於て，Shannon, Wiener が R.A. Fisher の導入した amount of information のある意味での拡張となる量を定義した。

Shannon-Wiener の amount of information とは，$x_1, \cdots, x_n$ を n 個の観測値，$f(x_1 \cdots x_n)$ をその確率分布密度函数とするとき

$$(1.1) \quad -\int \cdots \int \log f(x_1, \cdots, x_n; \theta_1^\circ, \cdots \theta_m^\circ) \cdot f(x_1, \cdots x_n; \theta_1^\circ \cdots \theta_m^\circ) dx_1 \cdots dx_n$$

により定義する。　茲に $\theta_1^\circ, \cdots \theta_m^\circ$ は補助変数である。

然し，これは，推定，検定に関係のない量である。

そこで，Bartlett は新しく information function

$$(1.2) \quad -\int \cdots \int \log f(x_1 \cdots x_n; \theta_1 \cdots \theta_m) \cdot f(x_1 \cdots x_n; \theta_1^\circ \cdots \theta_m^\circ) dx_1 \cdots dx_n$$

を導入した。　$\theta_1^\circ, \cdots, \theta_m^\circ$ は $\theta_1, \cdots, \theta_m$ と異なる量である。

(1.2) を用ひると R.A. Fisher の定義した amount of information と類似な性質をもつ。様な量を定義する事が可能である。　即ち，検定規準として用ひ、

$$(1.3) \qquad \frac{f(x_1, \cdots, x_n; \theta_1^0, \cdots, \theta_m^0)}{f(x_1 \cdots x_n; \theta_1, \cdots \theta_m)}$$

を考へる。　　　　　　　　　勿論同様である。から対数をとることにする。　$\theta_i = \theta_i^0$　$(i = 1, 2 \cdots m)$ なる仮定の下での数学的期待値

$$(1.4) \qquad I_f(\theta; \theta^0) = \int \cdots \int \log \frac{f(x_1 \cdots x_n; \theta_1^0 \cdots \theta_m^0)}{f(x_1, \cdots, x_n; \theta_1 \cdots \theta_m)} \cdot f(x_1 \cdots x_n; \theta_1 \cdots \theta_m) dx_1 \cdots dx_n$$

を考へる。　$I_f(\theta; \theta^0)$ は、伊藤氏による報知高と全く analogous な性質を有つ事がわかる。

筆者はそれに従って Bartlett の精しく述べてない事をきちんとやってみる事にした。

§ 2. $I_f(\theta; \theta^0)$ のいろいろな性質

<u>定理 1</u>　　$I_f(\theta; \theta^0) \geqq 0$

而も等号の成立するのは $\theta = \theta^0$ のときに限る。

<u>証明</u>　　$f(x_1 \cdots x_n; \theta_1, \cdots \theta_m) = g(x_1 \cdots, x_n)$　とおく。

そのとき

$$\int \cdots \int \log \frac{f}{g} \cdot f \cdot dx_1 \cdots dx_n$$

$0 \leqq \alpha \leqq 1$　に対して、$g_1 = (1-\alpha) f + \alpha g$　も又確率分布密度函数である。

$$-\int \cdots \int \log \frac{(1-\alpha) f + \alpha g}{f} \cdot f \, dx_1 \cdots dx_n$$

により、新しく $I'_f(\theta; \theta^0)$ を定義する。

Taylar 展開する事により

$$-\int\cdots\int \log\frac{(1-\alpha)f+\alpha g}{f}\cdot f\,dx_1\cdots dx_n$$

$$=-\int\cdots\int\left(0-\frac{(g-f)}{f}\alpha+\frac{\left(\frac{g-f}{f}\right)^2}{2\left(1+\frac{(g-f)}{f}\alpha^*\right)^2}\right)f\,dx_1\cdots dx_n$$

茲に α^* は $0\leqq\alpha^*=\alpha^*(x_1\cdots x_n)\leqq 1$ なる数である。
従って，特に $\alpha=1$ のとき，

$$-\int\cdots\int\log\frac{g}{f}\cdot f\,dx_1\cdots dx_n=-\int\cdots\int 0\,dx_1\cdots dx_n$$

$$-1+1+\int\cdots\int\frac{\frac{(g-f)^2}{(f)}}{2\left(1+\frac{(g-f)}{f}\alpha^{**}\right)^2}f\,dx_1\cdots dx_n$$

$$=\int\cdots\int\frac{\left(\frac{g-f}{f}\right)^2}{2\left(1+\frac{(g-f)}{f}\alpha^{**}\right)^2}f\,dx_1\cdots dx_n\geqq 0.$$

故に $I_f(\theta;\theta^0)\geqq 0$ となる。
Integrand は $f\neq 0$ のとき $g=f$ のとき 0 となり
$\int\cdots\int g\,dx_1\cdots dx_n=1$. $\int\cdots\int f\,dx_1\cdots dx_n=1$ なる條件
により，$\int\cdots\int(g-f)\,dx_1\cdots dx_n=0$ $g\neq f$ なる $(x_1\cdots x_n)$
の測度が正なるとき矛盾する。故に $g=f$（測度 0 を除く），
が成り立つ。従って定理を得る。

<u>定理 2</u> $(x_1,\cdots,x_n)$, $(y_1,\cdots y_n)$ が一対一且，補助変数に無関係に対応する時，

$$\int\cdots\int\log\frac{f(x_1\cdots x_n;\theta_1^0\cdots\theta_m^0)}{f(x_1\cdots x_n;\theta_1\cdots\theta_m)}\cdot f(x_1\cdots x_n;\theta_1^0\cdots\theta_m^0)\,dx_1\cdots dx_n$$

$$=\int\cdots\int\log\frac{g(y_1\cdots y_n;\theta_1^0\cdots\theta_m^0)}{g(y_1\cdots y_n;\theta_1\cdots\theta_m)}\cdot g(y_1\cdots y_n;\theta_1^0\cdots\theta_m^0)\,dy_1\cdots dy_n$$

茲に $g(y_1 \cdots y_n ; \theta_1^\circ, \cdots \theta_m^\circ)$ は $(y_1 \cdots y_n)$ の分布密度とする。

証　明

$$y_N = \varphi_N(x_1 \cdots\cdots x_n), \qquad (N = 1, 2 \cdots n)$$ と

する。 $\varphi_1, \cdots \varphi_n$ の函数行列式を Φ とする。

そのとき

$$g(\varphi_1(x_1 \cdots x_n), \varphi_2(x_1 \cdots x_n), \cdots \varphi_n(x_1 \cdots x_n); \theta_1 \cdots \theta_n)\, \Phi(x_1 \cdots x_n)$$
$$= f(x_1 \cdots x_n ; \theta_1, \cdots \theta_m)$$

$$\int \cdots \int \log \frac{f(x_1 \cdots x_n ; \theta_1^\circ \cdots \theta_m^\circ)}{f(x_1 \cdots x_n ; \theta_1 \cdots \theta_m)} \cdot f(x_1 \cdots x_n ; \theta_1^\circ \cdots \theta_m^\circ)\, dx_1 \cdots dx_n$$

$$= \int \cdots \int \log \frac{g(\varphi_1(x_1 \cdots x_n), \varphi_2(x_1 \cdots x_n) \cdots \varphi_n(x_1 \cdots x_n)) ; \theta_1^\circ \cdots \theta_m^\circ)}{g(\varphi_1(x_1 \cdots x_n)\, \varphi_2(x_1 \cdots x_n) \cdots \varphi_n(x_1 \cdots x_n) ; \theta_1 \cdots \theta_m)}$$
$$\times g(\varphi_1(x_1 \cdots x_n) \cdots \varphi_n(x_1 \cdots x_n)\, \theta_1^\circ \cdots \theta_m^\circ) \times \Phi(x_1 \cdots x_n)\, dx_1 \cdots dx_n$$

そこで $\Phi(x_1 \cdots x_n)$ は変換の Jacobian であるから、

$$\Phi(x_1 \cdots x_n)\, dx_1 \cdots dx_n = dy_1 \cdots dy_n$$

故に、$(x_1 \cdots x_n) \to (y_1 \cdots y_n)$ により

$$\int \cdots \int \log \frac{g(y_1 \cdots y_n ; \theta_1^\circ \cdots \theta_m^\circ)}{g(y_1 \cdots y_n ; \theta_1 \cdots \theta_m)} \cdot g(y_1 \cdots y_n ; \theta_1^\circ \cdots \theta_m^\circ)\, dy_1 \cdots dy_n$$

となる。 これより定理をうる。

$x_1 \cdots x_\ell$ の確率分布密度を $f(\sigma_1 \cdots \sigma_\ell ; \theta_1 \cdots \theta_m)$

$x_1 \cdots x_k$ の確率分布密度を $g(\sigma_1 \cdots \sigma_k ; \theta_1 \cdots \theta_m)$ とするとするとき

$$k(\sigma_1 \cdots \sigma_\ell ; \theta_1 \cdots \theta_m)\, d\sigma_{k+1} \cdots d\sigma_\ell = \frac{f(\sigma_1 \cdots \sigma_\ell ; \theta_1 \cdots \theta_m)}{g(\sigma_1 \cdots \sigma_k ; \theta_1 \cdots \theta_m)}\, d\sigma_{k+1} \cdots d\sigma_\ell$$

は $x_\kappa = \sigma_\kappa \quad K = 1.2 \cdots k$ なる條件の下での $x_{k+1} \cdots x_\ell$ の確率分布密度である。

そのとき次の定理が成り立つ。

定理 3.

$$I_f(\theta;\theta^\circ) = I_g(\theta;\theta^\circ) + \int\cdots\int I_h(\theta;\theta^\circ) g(\sigma_1\cdots\sigma_k;\theta_1^\circ\cdots\theta_m^\circ)\, d\sigma_1\cdots d\sigma_k$$

$$\geq I_g(\theta;\theta^\circ)$$

等号の成立つのは，$h(\sigma_1\cdots\sigma_\ell;\theta_1\cdots\theta_m)$ が θ に無関係なるときである。

証明

$m = 1$ のとき証明する。

$$\int\cdots\int \log\left(\frac{f(\sigma_1\cdots\sigma_\ell;\theta^\circ)}{f(\sigma_1\cdots\sigma_\ell;\theta)}\right) f(\sigma_1\cdots\sigma_\ell;\theta^\circ)\, d\sigma_1\cdots d\sigma_\ell$$

$$= \int\cdots\int \log\frac{g(\sigma_1\cdots\sigma_k;\theta^\circ)}{g(\sigma_1\cdots\sigma_k;\theta)} f(\sigma_1\cdots\sigma_\ell;\theta^\circ)\, d\sigma_1\cdots d\sigma_\ell$$

$$+ \int\cdots\int \log\frac{h(\sigma_1\cdots\sigma_\ell;\theta^\circ)}{h(\sigma_1\cdots\sigma_\ell;\theta)} f(\sigma_1\cdots\sigma_k;\theta^\circ)\, d\sigma_1\cdots d\sigma_k$$

$$= I_g(\theta;\theta^\circ) + \int\cdots\int g(\sigma_1\cdots\sigma_k;\theta^\circ)\, d\sigma_1\cdots d\sigma_k \int\cdots\int \log\left(\frac{h(\sigma_1\cdots\sigma_\ell;\theta^\circ)}{h(\sigma_1\cdots\sigma_\ell;\theta)}\right) h(\sigma_1\cdots\sigma_\ell;\theta^\circ)\, d\sigma_{k+1}\cdots d\sigma_\ell$$

$$\geq I_g(\theta;\theta^\circ)$$

となる。

玆に第二項が 0 となるのは，$h(\sigma_1\cdots\sigma_\ell;\theta^\circ)$ が θ に無関係なる時に限る事が定理 1 より分る。

従って定理をうる。定理の意味は報知高と同様である。

定理4.　$(x_1 ---- x_k)$ と $(x_{k+1}, ---, x_l)$ とが互に独立な観測値とする時

$$I_f(\theta;\theta^0) = I_g(\theta;\theta^0) + I_h(\theta;\theta^0)$$

が成り立つ。

証明　定理3より，明らかである。

§3. 充足統計量

$X_1, X_2, ----- X_n$ と n 個の確率変数とし，$t_1 ---- t_k$ をこれより導びかれた互に函数関係のない統計量とする。

且つその確率密度を $g(t_1 ---- t_k ; \theta_1 --- \theta_m)$ とする。

そのとき次の定理が成り立つ。

定理1

$$I_g(\theta;\theta^0) \leqq I_f(\theta;\theta^0)$$

特に等号の成り立つのは，$t_1, t_2, ---, t_k$ が定まった場合の $x_1 ---- x_n$ の分布法則が補助変数 $\theta_1, ---, \theta_m$ に無関係な事である。

証明　$m=1$ の場合に証明する事にする。

$t_1 ---- t_m$ に $t_{m+1} ---- t_n$ なる変数を補って $(t_1, ---- t_n)$ と $(x_1, ----, x_n)$ とが補助変数 $\theta_1, ---, \theta_m$ に無関係な対応で且一対一ならしめる。　そのとき前節の定理2，3により，

$$\int ----- \int \log \frac{f(x_1 ---- x_n ; \theta^0)}{f(x_1 --- x_n ; \theta)} f(x_1 --- x_n ; \theta^0) dx_1 --- dx_n$$

$$= \int ---- \int \log \frac{h(t_1 ---- t_n ; \theta^0)}{h(t_1 --- t_n ; \theta)} \cdot h(t_1 --- t_n ; \theta^0) dt_1 --- dt_n$$

$$\geqq \int \cdots \int \log \frac{g(t_1 \cdots t_m ; \theta^0)}{g(t_1 \cdots t_m ; \theta)} \cdot g(t_1 \cdots t_m ; \theta^0)\, dt_1 \cdots dt_m$$

玆で $t_1 \cdots t_n$ の確率分布密度函数を $h(t_1 \cdots t_n ; \theta^0)$ とする。

これで前半を生ずる §2. 定理3により等号の成り立つのは，$t_1 \cdots t_m$ が定まった場合の $t_1, t_2 \cdots t_n$ の分布法則が $\theta_1 \cdots \theta_m$ に無関係な場合であるが，$(t_1 \cdots t_n)$ と $(x_1 \cdots x_n)$ とが補助変数 $\theta_1 \cdots \theta_m$ に無関係な対応で一対一に対応する事から，$t_1 \cdots t_m$ が定まった場合に於ける $x_1 \cdots x_n$ の分布法法則が補助変数を含まない場合といふ事が出来る，

定理1より

$$0 \leqq \frac{I_g(\theta ; \theta^0)}{I_f(\theta ; \theta^0)} \leqq 1$$

が一般に成り立つ事がわかる。

而も，等号の成り立つのは充足統計量系の場合である事が次の定理により保証される。

<u>定理2.</u> $t_1, t_2, \cdots, t_m$ が $\theta_1 \cdots \theta_m$ の充足統計量系である為の必要且つ充分なる條件は，観測値 $x_1 \cdots x_n$ の分布密度を $f(\xi_1 \cdots \xi_n, \theta_1 \cdots \theta_m)$ が $f(x_1 \cdots x_n, \theta_1 \cdots \theta_m) = g(t_1 \cdots t_m, \theta_1 \cdots \theta_m) h(x_1 \cdots x_n)$ なる形になる事である。

従って，$I_f(\theta ; \theta^0) = I_g(\theta ; \theta^0)$ が成り立つのは f が上の様な函数形のときに限る事を示せばよい。

<u>証明.</u> (イ) 充分なる事 $(t_1 \cdots t_m)$ に $(t_{m+1} \cdots t_n)$ なる統計量を補って，$(t_1 \cdots t_m)$ と $(x_1 \cdots x_n)$ とが補助変数に無関係な対応で一対一に対応する様にする。

そのとき $(t_1 \cdots t_n)$, $(t_1 \cdots t_m)$ の分布密度を夫々 $p(t_1 \cdots t_n ; \theta_1 \cdots \theta_m)$, $k(t_1 \cdots t_m ; \theta_1, \cdots \theta_m)$ とする。

そこで $t_1 \cdots t_m$ が定まった場合の $t_1 \cdots t_n$ の分布法則が $\theta_1 \cdots \theta_m$ に無関係なる事を証明すればよい。

$$P_{t_1 \cdots t_m}(t_1 \cdots t_n ; \theta_1 \cdots \theta_m) = \frac{p(t_1 \cdots t_n ; \theta_1 \cdots \theta_m)}{R(t_1 \cdots t_m ; \theta_1 \cdots \theta_m)}$$ とする。

又初めに述べた事より

$$f(x_1 \cdots x_n ; \theta_1 \cdots \theta_m) = P(t_1 \cdots t_n ; \theta_1 \cdots \theta_m)\Phi(x_1 \cdots x_n)$$

$\Phi(x_1 \cdots x_n) = \dfrac{\partial(t_1 \cdots t_n)}{\partial(x_1 \cdots x_n)}$ で $\Phi(x_1 \cdots x_n)$ は $\theta_1 \cdots \theta_m$ に無関係である。

一方、 $t_\mu = \varphi_\mu(t_1 \cdots t_m)$ $\mu = 1, 2 \cdots m$ とする。

先づ分布密度の形が與えられた様であれば

$$g(t_1 \cdots t_m ; \theta_1 \cdots \theta_m) d\tau_1 \cdots d\tau_m = \underset{\substack{\tau_\mu < \varphi_\mu(\xi_1 \cdots \xi_n) < \tau_\mu + d\tau_\mu \\ \mu = 1, 2 \cdots m}}{\int \cdots \int} f(\xi_1 \cdots \xi_n ; \theta_1 \cdots \theta_m) d\xi_1 \cdots d\xi_m$$

$$= g(\tau_1 \cdots \tau_m ; \theta_1 \cdots \theta_m) \underset{\substack{(\tau_\mu < \varphi_\mu(\xi_1 \cdots \xi_n) < \tau_\mu + d_\mu) \\ \mu = 1, 2, \cdots m}}{\int \cdots \int} h(\xi \cdots \xi_n) d\xi_1 \cdots d\xi_n$$

$$= g(\tau_1 \cdots \tau_m ; \theta_1 \cdots \theta_m) S(\tau_1 \cdots \tau_m) d\tau_1 \cdots d\tau_m$$

故に

$$f(x_1 \cdots x_n ; \theta_1 \cdots \theta_m) = k(t_1 \cdots t_m, \theta_1 \cdots \theta_m) \frac{h(x_1 \cdots x_n)}{S(t_1 \cdots t_m)}$$

故に

$$P_{t_1 \cdots t_m}(t_1 \cdots t_n ; \theta_1 \cdots \theta_m) = \frac{S(t_1 \cdots t_m)}{h(x_1 \cdots x_n)\Phi(x_1 \cdots x_m)}$$

従って，$\theta_1 \cdots \theta_m$ を含まない事がわかる，

故に，

$$\int\cdots\int \log \frac{f(x_1\cdots x_n; \theta_1^0\cdots\theta_m^0)}{f(x_1\cdots x_n; \theta_1\cdots\theta_m)} f(x_1\cdots x_n; \theta_1^0\cdots\theta_m^0)dx_1\cdots dx_n$$

$$= \int\cdots\int \log \frac{k(t_1\cdots t_m; \theta_1^0\cdots\theta_m^0)}{k(t_1\cdots t_m; \theta_1\cdots\theta_m)} k(t_1\cdots t_m; \theta_1^0\cdots\theta_m^0)dt_1\cdots dt_m$$

が成り立つ。

（ロ）次に必要なる事を証明する。

今 $t_1 \cdots t_m$ が充足統計量系とする。 $t_1 \cdots t_m$ に $t_{m-1} \cdots t_n$ を附け加えて，$(x_1 \cdots x_n)$ と $(t_1 \cdots t_n)$ とが一対一に対応する様にする。 そうすれば，假定により，

$$I_f(\theta; \theta^0) = I_g(\theta; \theta^0)$$

である。 又 $(x_1 \cdots x_n)$ と $(t_1 \cdots t_n)$ とが一対一の対応をする事より，

$$I_f(\theta; \theta^0) = I_p(\theta; \theta^0)$$

但し，$(t_1 \cdots t_n)$，$(t_1 \cdots t_m)$ の分布密度を夫々，
$p(t_1 \cdots t_n; \theta_1 \cdots \theta_m)$ $g(t_1 \cdots t_m; \theta_1 \cdots \theta_m)$ とする。
$(x_1 \cdots x_n)$ と $(t_1 \cdots t_n)$ とが一対一に対応する事より，

$$f(x_1\cdots x_n; \theta_1\cdots\theta_m) = p(t_1\cdots t_n, \theta_1\cdots\theta_m)\Phi(x_1\cdots x_n)$$

$$\Phi(x_1\cdots x_n) = \frac{\partial(t_1\cdots t_n)}{\partial(x_1\cdots x_n)}$$

次に $t_1 = \tau_1 \cdots \tau_m = \tau_m$ といふ條件の下に於ける $(t_{m+1} \cdots t_m)$ の分布法則は $\theta_1 \cdots \theta_m$ に無関係である。（§3，定理1）
従って

$$\frac{p(t_1\cdots t_m; \theta_1\cdots\theta_m)}{g(\tau_1\cdots\tau_m; \theta_1\cdots\theta_m)} = \Phi(\tau_1\cdots\tau_m)$$

故に

$$f(x_1 \cdots x_n ; \theta_1 \cdots \theta_m) = g(t_1 \cdots t_m ; \theta_1 \cdots \theta_m)\, \Phi(x_1 \cdots x_n)\, \Phi(t_1 \cdots t_m)$$

$$= g(t_1 \cdots t_m ; \theta_1 \cdots \theta_m)\, h(x_1 \cdots x_n)$$

$$h(x_1 \cdots x_n) = \Phi(x_1 \cdots x_n)\, \Phi(t_1 \cdots t_m),$$ 証明終り

以上の事より $I_f(\theta ; \theta')$ を maximum にするものは充足統計量系に限る事が証明されたわけである。

II. 尤度比函数の積率母函数について。

1). 分布函数の集まり $\{F\} = \Omega$ に対して，次の様な假設をもうける。 Ωの各要素は測度mに関して絶対連続であるとする。 即ち

$$F(E) = \int_E f(x)\,dm$$

と表はされるとする。

假設 H_0 の下で Ω の一つの分布函数F_0を假設 H_1 の下で Ω の一つの分布函数 F_1 をとるものとする。

このとき，尤度比函数は $\dfrac{f_0(x)}{f_1(x)}$ により表はされる。

対数をとって，$\log \dfrac{f_0(x)}{f_1(x)}$ を考える。 その假設 H_0, H_1 の下での積率母函数を夫々 $M_0(t)$, $M_1(t)$ で表はす。 即ち，

$$M_0(t) = \int \left(\frac{f_0(x)}{f_1(x)}\right)^t f_0(x)\,dm$$

$$M_0(t) = \int \left(\frac{f_0(x)}{f_1(x)}\right)^t f_1(x)\,dm$$

である。 上の式より，$M_0(t) = M_1(t-1)$ が成り立つ事がわかる。

$M_1(t)$ に関して次の定理が成り立つ。

定理 1

$m\{x \mid f_0(x) \neq f_1(x)\} > 0$ なるとき　$M_1(t)$ は $0<t<1$ なるすべての t に対して存在し，而も，$M_1(t^*) < M_1(t)$ で $0<t^*<1$ なる t^* が一つ定まる。

証明　i). Minkowski の定理により，

$$M_1(t) = \int f_0(x)^t (f_1(x))^{1-t} dm \leqq \left(\int f_0(x)\right)^t \left(\int f_1(x)\right)^{(1-t)} = 1$$

ii)　$M_1(S) = \int f_0(x)^S f_1(x)^{1-S} dm$　とする。

$$M_1\left(\frac{S+t}{2}\right) \leqq \left(M_1(S) M_1(t)\right)^{\frac{1}{2}}$$

が Schwarz の不等式を用ひて証明される。

$M_1(S)$, $M_1(t)$ は非負の函数であるから $(M_1(S)M_1(t))^{\frac{1}{2}} \leqq \frac{M(S)+M_1(t)}{2}$

が成り立つ。　従って，$M_1\left(\frac{S+t}{2}\right) \leqq \frac{M_1(S)+M(t)}{2}$　が $0<S,t<1$

の任意の $S\cdot t$ に対して成り立つ。　故に，$M_1(S)$ は $0 \leqq t \leqq 1$ で下に凸の函数であることがわかる。

而も $M_1(t)$ は連続な函数なる事がわかるから $M_1(t^*) < M_1(t)$　なる $0<t^*<1$ が存在する。

$f_0(x) = f_1(x)$ a.e　なるとき，$M_1(t) = 1$　である。

ここで　$\inf M_1(t) = \inf \int f_0(x)^t (f_1(x))^{1-t} dm = M_1(t^*) = \rho(F_0, F_1)$

と定義すると上の定理により $m\{x \mid f_0(x) \neq f_1(x)\} > 0$ のとき

$\rho < 1$ が成り立ち　$\rho = 1$ が成り立つのは

$$m\{x \mid f_0(x) = f_1(x)\} = 1$$

のときである。　この $\rho(F_0, F_1)$ は仮説の間のある意味での測度になってゐる。これは，Herman chernoff (1) による尤度比検定の index と呼ばれるものである。この ρ は松下嘉米男氏の affinity と類似な量である。　　　　従って，全く同様にして次の定理

が成り立つ。

定　理　2

ρを上に定義した数とする。　εを任意の正数とするとき，適当に $k(\varepsilon)$ をとり，$\rho^k<\varepsilon$ ならしめる。

今

$$E_0=\{(x_1,x_2,\cdots,x_k);\ f_1(x_1)\cdots f_1(x_k)\geqq f_2(x_1)\cdots f_2(x_k)\}$$

を標空間 k 一次元ユークリッド空間上の点集合とする。

$E_1=E_0^c$ とする。　そのとき，

$$F_P^{(k)}(E_1)>1-\varepsilon,\qquad F_1^{(k)}(E_2)>1-\varepsilon$$

が成り立つ。　茲で $F_0^{(k)}$ $F_1^{(k)}$ は夫々 F_0 F_1 の R^k 上の分布函数である。

証　明

$$F_0^{(k)}(E_0)=\int_{E_0} f_0(x_1)\cdots f_0(x_k)dm\cdots dm=1-\int_{E_2} f_0(x_1)\cdots f_0(x_k)dm\cdots dm$$

$$\geqq 1-\int_{E_1}(f_0(x_1)\cdots f_0(x_k))^{t^*}(f_1(x_1)\cdots f_1(x_k))^{1-t^*}dm\cdots dm$$

$$=1-\left(\int_{E_1} f_0(x)^{t^*}f_1(x)^{1-t^*}dm\right)^k$$

$$=1-\rho^k>1-\varepsilon$$

$F_1^{(k)}(E_1)>1-\varepsilon$　についても全く同様にして証明される。

上の定理により，標本$(x_1\cdots x_k)$が E_0 にあるとき，F_0，E_0 にあるとき F_1 をとる事にすれば，$(x_1\cdots x_k)$が F_0 からの標本でないとき F_1 であるとする過誤，F_1 の標本であるとき，F_0であるとする過誤を夫々εより少さくしうる事が示される。

（参照 Kameo Matsusita(3)）

2)　(2) によれば $t=\frac{1}{2}$ としてあるけれども，実は $t=\frac{1}{2}$ とする事が此の様な立場からでも実際に非常によい事が多くの場合に示される。

例1.　母分散が等しく，平均のみ異る正規分布の場合，

$$f_0(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2\sigma^2}(x-m_0)^2},\quad f_2(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2\sigma^2}(x-m_1)^2}$$

とするとき，

$$\rho(F_0,F_1)=\inf_{0<t<1}\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)e^{-\frac{1}{2\sigma^2}t(x-m_0)^2-\frac{1}{2\sigma^2}(x-m_1)^2(t-1)}dx=e^{-\frac{1}{8\sigma^2}(m_0-m_1)^2}$$

となり $t=\frac{1}{2}$ とする時と一致する。

例2.　補助変数の異なるポアソン分布の場合。

$$f_0(x)=\frac{\eta_0^x}{x!}e^{-\eta_0}\qquad f_1(x)=\frac{\eta_1^x}{x!}e^{-\eta_1}\qquad (x=0,1,2,----)$$

とするとき，

$$\rho(F_0,F_1)=\inf_{0<t<1}\sum_{x=0}^{\infty}\frac{\eta_0^{xt}e^{-\eta_0 t}\eta_1^{x(1-t)}e^{-\eta_1(1-t)}}{x!}$$

$$=\inf_{0<t<1}e^{-\eta_0 t-\eta_1(1-t)}e^{+\eta_0 t-\eta_1^{(1-t)}}$$

t の値は次の式を満足する。

$$-\eta_0+\eta_1+\eta_0^t\eta_1^{(1-t)}(\log\eta_0-\log\eta_1)=0$$

$\frac{\eta_0}{\eta_1}=\alpha$　とするとき

$$t=\frac{\log\frac{1-\alpha}{\log\frac{1}{\alpha}}}{\log\alpha}=\frac{\log\frac{\log\frac{1}{\alpha}}{1-\alpha}}{\log\frac{1}{\alpha}}$$

$\alpha\to 0$ のとき $t\to 1$　$\alpha\to 1$ のとき $t=\frac{1}{2}$ である。

例へば　$\alpha=\frac{1}{2}$　のとき，

$$t = 1+\frac{\log\log 2}{\log 2} = 1-\frac{0.159}{0.301} \fallingdotseq 0.47$$

となり $\frac{1}{2}$ に非常に近い事がわかる．

例3．　補助変数の一つのみが異る一般の χ^2－分布に対して，

$$f_0(x)=\frac{\alpha_0^{\lambda}}{\Gamma(\lambda)}x^{\lambda-1}e^{-\alpha_0 x} \qquad f_1(x)=\frac{\alpha^{\lambda}}{\Gamma(\lambda)}x^{\lambda-1}e^{-\alpha x}$$

とするとき，

$$\rho(F_0,F_1)=\inf_{0<t<1}\int_0^{\infty}\frac{x^{\lambda-1}}{\Gamma(\lambda)}e^{-(\alpha_0 t+\alpha_1(1-t))x}\alpha_0^{\lambda t}\alpha_1^{\lambda(1-t)}dx$$

$$=\inf_{0<t<1}\left(\frac{\alpha_0^{t}\alpha_1^{(1-t)}}{\alpha_0 t+\alpha_1(1-t)}\right)^{\lambda}$$

これを満足する t は Poisson－分布の場合と同様である。

あとがき．

前半については，Kullback Leiler((2)参照)更に一般的に論じてゐる事を知った。　けれども一応もとのまゝで出す事にいたしました。　此の小文を書くにあたって，随分御面倒を鍋谷さんにおかけしました。　御礼を申し上げます。

参 考 文 献

(1) M. S. Bartlett. The Statislical approach to the analysis of Time-series

(2) 伊藤 清. 報知高 統計数理研究所 講究録

(3) C. E. Shannon The mathematical theory of Communication

(4) N. Wiener. Cyhemetics

後半については

(1) Herman Chernoff. A measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations. (Annals of Mathematical Statistics Vol. 28. No. 4).

(2) S. Kullback. An application of Information Theory to Multivariate Analysis (Annals of Mathematical Statistics Vol. 23. No. 1)

(3) Kameo Matsushita and Hirotugu Akaike. Note on the decision problem (Annals of Institute of Statistical mathematics Vol. IV. No. 1)