数理解析研究所講究録
590巻 1986年 20-32

仮説に制約条件がある場合の Bivariate Sign Test

九大理学部　野間口謙太郎

Hodges (1955 AMS) は1次元 data に対する sign 検定を uion-intersection principle を用いて、2次元 dataに対して拡張した。これがいわゆる bivariate sign testである。彼の考え方は parameter space に cone-制約条件が付いた場合でも適用可能である。

R^2上の分布関数 F を,原点を境界に持つ任意の half space 上の確率が常に1/2 であるものとする。C を R^2 での真の closed convex cone とする。一般性を失わずに一つの edge を x -軸に取ってよい。さらにもう一つの edge を vector L 方向とする。ここで vector の argument 関数 arg(・) を通常の argument 関数を π で割ったものとする。例えば arg((0,1)')は1/2となる。これを用いて C は $\{x \in R^2 \mid 0 \leq \arg(x) \leq \arg(L)\}$ とかける。このとき sample X1, X2, ・・・, Xn ～ F (x - θ) に基づいて、

仮説検定問題 $H_0: \theta = 0$ vs. $H_1: \theta \in C - \{0\}$ に対する、Hodges 流の検定統計量 T_n を提案したい。ここでの取り扱い方は Joffe & klotz (1962 AMS) に従っていて、この統計量の exact 分布、漸近分布を求め、 Bahadur efficiency を計算する。

まず Hodges が、どのように考えたか、みてみたい。彼は

(1) 任意の $\mu \in R^2$ に対して $H_0(\mu) \equiv H_0$、$H_1(\mu) \equiv \{\mu\}$ と置き、仮説を次のように分解した。

$$H_0 = \bigcap_{\mu \in R^2} H_0(\mu) \quad , \quad H1 = \bigcup_{\mu \in R^2} H_1(\mu)$$

(2) 次ぎに、部分仮説検定問題 $H_0(\mu)$ vs. $H_1(\mu)$ に対する検定統計量として $\mu' X1, \cdots, \mu' Xn$ を用いた sign 検定を採用する。つまり

$$T_\mu \equiv (1/n) \sum_{i=1}^{n} I\ [\mu' X i > 0]$$

を用いる。

(3) union-intersection principle の考え方により本来の検定統計量として次の

$$T \equiv \sup_{\mu \neq 0} T_\mu$$

を採用する。 T が大きいとき H_0 が棄却されることになる。

これが、Hodges のとった approach である。これにならって制約条件付の場合も同様にしたいのであるが、よくながめると（2）において必ずしも X i と対立仮説の方向 μ との内積をとる必要はないということが解る。内積をとる相手としてはもっと良い方向 ν がありうる。次の例を考えてみる。各 X i が、elliptical 分布に従っているとする。つまり $X_i = V_i \Sigma^{1/2} U_i$ という形をしているとしよう。ただし、Σ：positive definite matrix、U_i：単位円周上の一様分布に従う r.v.、V_i：正の r.v. とする。このとき、$H_0(\mu)$ vs. $H_1(\mu)$ に対する sign 検定の中で most powerful な方向は、

$$P(\nu' X_i \geq 0 \mid \theta = \mu)$$
$$= P(\nu'(V_i \Sigma^{1/2} U_i + \mu) \geq 0)$$
$$= E(P(\nu' \Sigma^{1/2} U_i \geq V_i^{-1} \nu' \mu \mid V_i))$$

を最大にするものであるから、$\nu = \Sigma^{-1} \mu$ とすれば良いことがわかる。勿論 Σ は未知であるから直接にこれとの内積はとれないのであるが、$\Sigma = I$ の場合は μ との内積をとれば良いのだから、data X_i をなるべく spherical 分布に従うように変換した後で対立仮説方向との内積を取ることにすれば良い。この考え方により次ぎの検定統計量を提案する。

（1） 各 data の原点からの距離を一定にする。

$$Y_i \equiv X_i / | X_i |$$

（２）　第３、4象限の data は原点に関して対称な点に移す。

$$Z_i \equiv \begin{cases} Y_i & \text{if } 0 \leq \arg(Y_i) < \pi \\ -Y_i & \text{otherwise} \end{cases}$$

（３）　Z_iをその argument の順序を守りながら第１象限の半円周上に等間隔に並べる。Z_iに対応するものをU_iとおく。つまりZ_iの argument が j番目ならば $\arg(U_i) = (j/n) - (1/2n)$であるようにとる。　このとき $\rho \equiv P(X_i \in C \cup (-C))$、この一致推定量を

$$\hat{\rho} \equiv (1/n) \sum_{i=1}^{n} I [Z_i \in C]$$

とおき，vector L’，cone C’を次のように定義する。

$$\arg(L') = \hat{\rho},$$

$$C' \equiv \{ x \mid 0 \leq \arg(x) \leq \arg(L') \}。$$

L’とC’は、Z_iからU_iへの変換に伴うLとCの変換に対応する。　任意の$\nu \in R^2$に対しても同様に、$\nu' \in R^2$を定義する。

（４）（２）において対称な点に移したものは再度対称な点に返す。

$$V_i \equiv \begin{cases} U_i & \text{if } Z_i = Y_i \\ -U_i & \text{otherwise} \end{cases}$$

(５)　$R_t \equiv \{X \in R^2 \mid -1/2 < (\arg(X)-t) \bmod 2 \leq 1/2\}$

と置き、

$$T_n(t) \equiv \sum_{i=1}^{n} I\,[V_i \in R_t]$$

を定義して、検定統計量として

$$T_n \equiv \sup_{0 \leq t \leq \rho} T_n(t)$$

を採用する。 α-point を $k_\alpha(\rho)$と置くと、検定は

$$T_n \geqq n/2 + \sqrt{n}\,k_\alpha(\hat{\rho})$$

のとき、H_0を棄却する事になる。

なお、制約条件が無い場合には optimal な方向などというようなものを考える必要はない。 なぜならば、全ての方向との内積をとってその sup をみるのだから、結局 optimal な方向との内積も考慮されていることになるからである。

＊＊＊　ＥＸＡＣＴ　分布　の　計算　＊＊＊

まず、$C = \{x \in R^2 \mid 0 \leq \arg(X) \leq 1\}$ の場合を考えてみる。 この場合の統計量を特別に T_n^* と書くことにする。 $T_n(i/n) - T_n(0)$, $i=0,1,2,\cdots,n$ が random walk であることは容易に確かめることができるので、Si，i

=0,1,2,・・・,n を random walk とするとき

$$P(2(T_n^* - n/2) \leq t)$$

$$= P(2 \sup_{0 \leq i \leq n} S_i - S_n \leq t)$$

と書ける. ところで、鏡像原理によると

$$P(S_1 < a, \cdots, S_n < a, S_n = b)$$

$$= P(S_n = b) - P(S_n = 2a - b)$$

であるから、n, t が偶数のとき、

$$P(2(T_n^* - n/2) \leq t)$$

$$= \sum_{i=-t}^{t} P(S_j < (t+i)/2, j=1,\cdots,n, S_n = i)$$

$$= P(-t \leq S_n \leq t) - (t+1)P(S_n = t+2)$$

n, t が奇数の場合も同様にして、

$$= P(-t \leq S_n \leq t) - tP(S_n = t+2)$$

を得る。

一般の $C = \{X \in R^2 \mid 0 \leq \arg(X) \leq \arg(L)\}$ の場合は、$C \cup (-C)$ に落ちた data の数 N_0 の条件付のもとで求められる。 明らかに $N_0 \sim B(n, \rho)$ であるが、$N_0 = n_0$ が与えられたとき、$\bigcap_{0 \leq t \leq \arg(L')} R_t$ に含まれる data U_i は常に T_nに +1として貢献 するので、その数を Y_{n-n_0}と置く、$R^2 - C \cup (-C)$ に含まれる n_0個の U_iに関

しては $C=\{x\in R^2 \mid 0 \leqq \arg(X) \leqq 1\}$ の場合に考察したものと同様であるから $T^*_{n_0}$ と書く。 すると $T_n = Y_{n-n_0} + T^*_{n_0}$ と表現でき、$N_0 = n_0$ の条件付で、 $Y_{n-n_0} \perp\!\!\!\perp T^*_{n_0}$ 、Y_{n-n_0} である。 これより T_n の分布を求めるには、N_0 が与えられたもとでの Y_{n-n_0} と $T^*_{n_0}$ との convolution をとり、しかるのちに N_0 の分布で mixture をとらなければならない。 この計算は不可能とは言わないが、大変にうるさい。
そこで、次ぎに漸近分布の計算をやってみる。

＊＊＊ 漸近分布 の 計算 ＊＊＊

まず T^*_n の漸近分布を求めてみる。 exact 分布より
n が偶数のとき、

$$
\begin{aligned}
P(2(T^*_n - n/2) \leqq \sqrt{n}\,t)
&= P(-[\sqrt{n}\,t] \leqq S_n \leqq [\sqrt{n}\,t]) \\
&\quad - ([\sqrt{n}\,t]+1)P(S_n = [\sqrt{n}\,t]+2)
\end{aligned}
$$

である。 ただし $[\sqrt{n}\,t]$ は $\sqrt{n}\,t$ より小さい最大の偶数である。
またnが奇数の場合も同様であるから、これよりnについて極限をとって、次を得る。

$$\lim_{n\to\infty} P(2(T^*_n - n/2) \leqq \sqrt{n}\,t)$$

$$= \Phi(t) - \Phi(-t) - t\,\phi(t)$$

ただしΦ、φ はそれぞれ標準正規分布 N(0,1)の分布関数と密度関数である。この極限分布は、Zを自由度3のカイ二乗分布に従うとするときの$Z^{1/2}$ の従う分布に等しいので、これを $\sqrt{\chi^2_3}$ 、この分布関数を G と書くことにする。また $\sqrt{\rho}$ Zの分布を $\sqrt{\rho}\sqrt{\chi^2_3}$、N(0,1-ρ)を $\sqrt{1-\rho}$ N(0,1)、これらの convolution を $\sqrt{1-\rho}$ N(0,1) + $\sqrt{\rho}\sqrt{\chi^2_3}$、この分布関数を G_ρ と書くことにする。 exact 分布の表現から、

$$2(T_n - n/2)/\sqrt{n} \underset{a}{\sim} \sqrt{1-\rho}\, N(0,1) + \sqrt{\rho}\sqrt{\chi^2_3}$$

となるのはすぐに解るのであるが、ここでは別の approach を試みてみよう。

$$(T_n - n/2)/\sqrt{n}$$

$$= \sup_{0\le t\le \hat{\rho}} (T_n(t) - T_n(0))/\sqrt{n} - (T_n(1) - T_n(0))/2\sqrt{n}$$

と変形できるが、$\{(T_n(t) - T_n(0))/\sqrt{n},\ 0 \leqq t \leqq 1\}$ は process として Wiener process W_t に、$\hat{\rho}$ は ρ に弱収束するので、これより

$$2(T_n - n/2)/\sqrt{n} \underset{a}{\sim} 2\sup_{0\le t\le \rho} W_t - W_1$$

$$= (2\sup_{0\le t\le \rho} W_t - W_\rho) + (W_\rho - W_1)$$

を得る。 明らかに第1項と第2項とは独立であり、第2項

は $\sqrt{1-\rho}$ N（0,1）に従う。　よって、第1項が $\sqrt{\rho}\sqrt{\chi_3^2}$ に従うことを示せばよい。　このために $B_t \equiv W_{t\rho}/\sqrt{\rho}$ ，$0 \leq t \leq 1$，$(M, B) \equiv (\sup_{0\leq t\leq 1} B_t, B_1)$ と定義する。

B_tはまた Wiener process となり、この時の（M，B）の同時密度関数が次ぎで与えられる事はよく知られている。

$$f(m,b)=\begin{cases}\sqrt{2/\pi}\,(2m-b)\exp(-(2m-b)^2/2) & \\ \qquad\qquad\qquad \text{if} \quad m\geqq 0, m\geqq b \\ 0 \qquad\qquad\qquad \text{otherwise}\end{cases}$$

これより $2M-B = (2\sup_{0\leq t\leq \rho} W_t - W_\rho)/\sqrt{\rho}$ の密度関数は

$$\sqrt{2/\pi}\ t^2 \exp(-t^2/2)$$

と計算できる。　これは $\sqrt{\chi_3^2}$ の密度関数である。

＊＊　Approximate Bahadur Effciency の 計算　＊＊

検定統計量 Sn の Bahadur slope は、

（1）　$\theta=0$ の下で、Sn はある確率変数Sに弱収束し、

Sの分布関数をF_Sと置くとき、

$\log(1-F_S(x)) = (-a\ x^2/2)(1+o(1))$

となり、

（２） $\theta=\mu$ の下で、$S_n/\sqrt{n}$ が $b(\mu)$ に確率収束する

とき、

$a(b(\mu))^2$ で与えられる。

ここでは $S_n=(2/\sqrt{n})(T_n-n/2)$ と置く。 さらに $\theta=0$ の下での分布を parameter（Σ、0）の elliptical 分布として、$\Sigma=I$ のときの X－座標の周辺分布を Ψ で表すことにする。

a の 計算

Tn の定義から明らかに、

$$T_n(0) \leqq T_n \leqq T_n^*$$

である。よって

$$(2/\sqrt{n})(T_n(0)-n/2) \leqq S_n$$
$$\leqq (2/\sqrt{n})(T_n^*-n/2)$$

これより次ぎの stochastic larger 性 を得る。

$$N(0,1) \underset{St.}{<} \sqrt{1-\rho}\,N(0,1)+\sqrt{\rho}\sqrt{\chi_3^2} \underset{St.}{<} \sqrt{\chi_3^2}$$

これは分布関数で書けば

$$\Phi(x) \geqq G_\rho(x) \geqq G(x)$$

となるが、

$$\log(1-\Phi(x)))=(-x^2/2)(1+o(1))$$

$$\log(1-G(x)))=(-x^2/2)(1+o(1))$$

が成り立つのは明らかだろうから、次がえられる。

$$\log(1-G_\rho(x)))=(-x^2/2)(1+o(1))$$

つまり a=1 である。

b(μ) の 計算

これを求めるために、$\nu'\mu>0$ であり、$\nu' Z_i>0$ となる Z_i の個数が丁度 [n／2] 個となる様な方向 ν で、この長さが $|\Sigma^{-1}\mu|$ に等しくなるものを ν_n とおく。$\mu\in C$ を変換した $\mu'\in C'$ に対する $R_{\mu'}$ を考えると、この内部に含まれる V_i と、$\nu' Z_i>0$ となる Z_i とは互いに1対1に対応するので、

$$W_n(\nu)\equiv(1/n)\sum_{i=1}^{n} I[\nu' X_i>0]$$

と置くと、 $$W_n(\nu_n)\leqq(1/n)T_n$$

であるという事が解る。$\Sigma^{-1}\mu$ に直交する vector で上半平面にあるものを τ とするとき、領域 $\{x \mid \arg(\mu)\leqq\arg(x)\leqq\arg(\tau) \text{ or } \arg(-\mu)\leqq\arg(x)\leqq\arg(-\tau)\}$ の確率は elliptical 性から 1/2 に等しいので、定義より ν_n は $\Sigma^{-1}\mu$ の一致推定量になっていることも解る。また

$$(1/n)T_n\leqq(1/n)\sup_{0\leq t\leq 2} T_n(t)$$

$$=\sup_{\nu\neq 0} W_n(\nu)$$

である。 $W_n(\nu)$ は、ν の argument の関数としてながめるとき 定数関数 $P(\nu' X_i > 0 \mid \theta = \mu) = \Psi(\nu'\mu / (\nu'\Sigma\nu)^{1/2})$ に弱収束する。 このことより

$$\Psi(\Delta) \leqq \lim_{n\to\infty} (1/n) T_n$$

$$\leqq \sup_{\nu\neq 0} \Psi(\nu'\mu / (\nu'\Sigma\nu)^{1/2})$$

$$= \Psi(\Delta)$$

を得る。 ただし $\Delta = (\mu'\Sigma^{-1}\mu)^{1/2}$ である。 ゆえに

$$b(\mu) = \lim_{n\to\infty} (1/\sqrt{n})\ S_n$$

$$= \lim_{n\to\infty} (2/n)(T_n - n/2)$$

$$= 2\Psi(\Delta) - 1$$

となる。

以上により S_n の Bahadur slope は $(2\Psi(\Delta) - 1)^2$ となることが解る。これは結局、制約条件がない場合の Bahadur slope と同じである。 よって、この量によっては、制約条件がある場合と、ない場合との差はみつけられない。 また Hotelling の T^2 検定の slope は、分散が存在して Σ に等しいとき、Δ^2 になるので、我々の検定の T^2 検定に対する Bhadur efficiency は

$$(2\Psi(\Delta) - 1)^2 / \Delta^2$$

で得られる。これは1次元で　$\Delta=\mu/\sigma$　と置いたときの t 検定に対する sign 検定の Bhadur efficiency に等しい。$\Psi=\Phi$ で $\mu\to\infty$ としたとき、この値がその時の Pitman efficiency $2/\pi$ になるという事実も良く知られている。