とりあえず，この節というか１６章は６章の Contiguity と７章の Local Asymptotic Normality を読んでいないと全く読めないようなので，必要な範囲で６，７章の記号，結果を説明しておく．

6 Contiguity

”Contiguity” = ”asymptotic absolute continuity”である．Contiguity は分布 P_n の極限分布から Q_n に従う統計量の列の極限分布を求めるためのテクニックを扱う．特に，P_n が帰無分布で Q_n が対立仮説下の分布である場合などに興味が有る．

6.1 Liklihood Ratios

まず, 記号の定義．$(\Omega, \mathcal{A})$ 上の測度 P, Q を考える． Q が P に対して絶対連続とは任意の $A \in \mathcal{A}$ で, $P(A) = 0$ ならば $Q(A) = 0$ となること．$Q << P$ とかく．さらに P と Q が直交するとは，$\Omega = \Omega_p \cup \Omega_Q, \Omega_p \cap \Omega_Q = \emptyset, P(\Omega_Q) = 0 = Q(\Omega_P)$ となることをいう．$P \perp Q$ とかく．

P と Q は測度 μ に対する密度 p と q を持つとし，

$$\Omega_P := \{p > 0\}, \;\; \Omega_Q := \{q > 0\}$$

と定義する．つまりそれぞれ，P と Q のサポートである．

測度 Q を $Q = Q^a + Q^{\perp}$ と測度の和で書く．ただし，

$$Q^a(A) = Q(A \cap \{p > 0\}); \;\; Q^{\perp}(A) = Q(A \cap \{p = 0\}).$$

明らかに $Q^a << P$ かつ $Q^{\perp} \perp P$ である．さらに任意の $A \in \mathcal{A}$ に対して，

$$Q^a(A) = \int_A \frac{q}{p} dP$$

が成り立つことが容易に分かる．分解 $Q = Q^a + Q^{\perp}$ は Q の P に対するルベーグ分解と呼ばれる．Q^a と $Q^{\perp}$ はそれぞれ Q の P に対する absolutely continuous part と othogonal part と呼ばれる．関数 q/p は Q^a の P に対する密度である．

$$\frac{dQ}{dP} = \frac{q}{p}, \;\; P\text{- a.s.}$$

とかく（dQ^a/dP でないことに注意）．ここからが本題であるが，今回の１６章の範囲で用いられる記号としては，これで十分なので７章に進む

7 Local Asymptotic Normality

7.1 Using Local Asymptotic Normality

まず，局所漸近正規性とは，尤度比の過程が漸近的に，正規分布の位置パラメータに対する尤度比過程と似ていることをいう．

このことをもう少し正確に定義していく．

$X_1, ..., X_n \sim P_\theta$, P_θ はある可測空間 $(\mathcal{X}, \mathcal{A})$ 上の $\theta \in \Theta \subset \mathbb{R}^k$ でインデックスされた分布とする．full observation は P_θ^n からの１つの観測と考える．すると統計モデルは $(\mathcal{X}^n, \mathcal{A}^n)$ 上の確率測度の族 $\{P_\theta^n : \theta \in \Theta\}$ で記述できる．

ある既知の固定されたパラメータ θ_0 周りで中心化し（reparametrization），局所パラメータ $h = \sqrt{n}(\theta - \theta_0)$ を定義することで，P_θ^n は $P^n_{\theta_0 + h/\sqrt{n}}$ と書ける．このとき，オリジナルの統計的実験 $\theta \mapsto P_\theta$ がパラメータ上で"smooth"であるなら，大きな n に対し，２つの統計的実験

$$(P^n_{\theta_0 + h/\sqrt{n}}) \;\text{ and }\; (N(h, I_{\theta_0}^{-1}) : h \in \mathbb{R}^k)$$

が似通った統計的性質を持っていることが示される．２つ目はシンプルな実験で，容易に解析できる．

7.2 節 Expanding the Likelihood では，局所実験の収束が定義される．以下の Theorem が局所漸近正規性を定義する．

Theorem 7.2. 開集合 $\Theta \subset \mathbb{R}^k$ と θ で differentiable in quadratic mean なモデル $(P_\theta : \theta \in \Theta)$ を考える．すると，$P_\theta \dot{\ell}_\theta = 0$ で，フィッシャー情報行列 $I_\theta = P_\theta \dot{\ell}_\theta \dot{\ell}_\theta^T$ が存在する．さらに，任意の $h_n \to h (n \to \infty)$ に対して，

$$\log \prod_{i=1}^n \frac{p_{\theta + h_n/\sqrt{n}}}{p_\theta}(X_i) = \frac{1}{\sqrt{n}} \sum_{i=1}^n h^T \dot{\ell}_\theta(X_i) - \frac{1}{2} h^T I_\theta h + o_{p_\theta}(1).$$

ただし，モデル $(P_\theta : \theta \in \Theta)$ が differentiable in quadratic mean at θ であるとは，ある $\dot{\ell}_\theta = (\dot{\ell}_{\theta,1}, ..., \dot{\ell}_{\theta,k})^T$ が存在して，

$$\int \left[\sqrt{p_{\theta+h}} - \sqrt{p_\theta} - \frac{1}{2} h^T \dot{\ell}_\theta \sqrt{p_\theta} \right]^2 d\mu = o(||h||^2), \quad h \to 0$$

が成り立つことをいう．

16.3 Using Local Asymptotic Normality

帰無仮説 $H_0 : \theta \in \Theta_0$，対立仮説 $H_1 : \theta \in \Theta_1$ の検定で，$\Theta = \Theta_0 \cap \Theta_1$ とおくと，尤度比統計量は

$$\Lambda_n = 2 \log \frac{\sup_{\theta \in \Theta} \Pi_{i=1}^n p_\theta(X_i)}{\sup_{\theta \in \Theta_0} \Pi_{i=1}^n p_\theta(X_i)}$$

で定義された．

尤度比統計量の漸近分布は実験の収束に基づく．このアプローチは一般の実験において可能であるが，この章では局所漸近正規性に制限する．このアプローチは線形でないパラメータ空間にも適用できる．

局所パラメータ空間 $H_n = \sqrt{n}(\Theta - \vartheta)$ と $H_{n,0} = \sqrt{n}(\Theta_0 - \vartheta)$ を導入する．尤度比統計量は

$$\Lambda_n = 2 \sup_{h \in H_n} \log \frac{\Pi_{i=1}^n p_{\vartheta + h/\sqrt{n}}(X_i)}{\Pi_{i=1}^n p_\vartheta(X_i)} - 2 \sup_{h \in H_{n,\vartheta}} \log \frac{\Pi_{i=1}^n p_{\vartheta + h/\sqrt{n}}(X_i)}{\Pi_{i=1}^n p_\vartheta(X_i)}$$

という形で書ける．

Chapter 7 では，大きな n に対して上式のようなリスケールされた尤度比過程は，正規分布の実験の尤度比過程 $(N(h, I_\theta^{-1}) : h \in \mathbb{R}^k)$ に似ていることが示された．このことは，もし集合 H_n と $H_{n,0}$ が適切な意味で集合 H と H_0 に収束するならば，列 Λ_n は

$$\Lambda = 2 \sup_{h \in H} \log \frac{dN(h, I_\vartheta^{-1})}{dN(0, I_\vartheta^{-1})}(X) - 2 \sup_{h \in H_\vartheta} \log \frac{dN(h, I_\vartheta^{-1})}{dN(0, I_\vartheta^{-1})}(X)$$

で与えられる代替的な正規尤度比により得られる確率変数 Λ に分布収束することを意味している．

上式はまさに観測 X に基づく帰無仮説 $H_0 : h \in H_0$ vs 対立仮説 $H_1 : h \in H - H_0$ の仮説検定に対する尤度比統計量である．後者の実験はシンプルであるので，このヒューリスティックは列 Λ_n の漸近分布を得るためだけではなく，対応する検定の列の漸近的な質を理解するのに役立つ．

正規実験に対する尤度比検定統計量は 7 章（p.101）の結果

$$\log \frac{dN(h, I_\theta^{-1})}{dN(0, I_\theta^{-1})}(X) = -\frac{1}{2}(X-h)^T I_\theta (X-h) + \frac{1}{2} X^T I_\theta (X)$$

を用いれば（これは容易に分かる），

$$\begin{aligned}\Lambda &= \inf_{h \in H_0} (X-h)^T I_\vartheta (X-h) - \inf_{h \in H} (X-h)^T I_\vartheta (X-h) \\ &= ||I_\vartheta^{1/2} X - I_\vartheta^{1/2} H_0||^2 - ||I_\vartheta^{1/2} X - I_\vartheta^{1/2} H||^2\end{aligned}$$

となる．ただし，ノルムは今回のように元が集合のときは，その集合内の最近傍までの距離とする．

ϑ のもとでの列 Λ_n の分布は $h=0$ のもとでの λ の分布に一致することが分かる．$h=0$ のもとでベクトル $I_\vartheta^{1/2} X$ は標準正規分布に従う．以下の補題は線形部分空間への標準正規確率変数の 2 乗距離がカイ二乗分布に従うことを示し，H_0 が線形空間のときの尤度比統計量のカイ二乗極限を説明する．

Lemma 16.6. X を k 次元の標準正規分布に従う確率変数とし，H_0 は $\mathbb{R}^k$ の ℓ 次元線形部分空間とする．このとき，$||X - H_0||^2$ は自由度 $k-\ell$ のカイ二乗分布に従う．

Proof. $\mathbb{R}^k$ の直交基底を最初の ℓ 個が張る空間が H_0 となるようにとる．ピタゴラスの定理より，ベクトル z から H_0 への距離の 2 乗は，この基底への $k-\ell$ 個の成分の 2 乗和 $\sum_{i>\ell} z_i^2$ に等しい．

基底の変換は成分の直交変換に等しい．標準正規分布は直交変換において不変であるので，X の任意の基底の成分は独立な標準正規分布である．

よって，$||X - H_0||^2 = \sum_{i>\ell} X_i^2$ は自由度 $k-\ell$ のカイ二乗分布に従う． □

もし，ϑ が Θ の内点であれば，集合 H は $\mathbb{R}^k$ 全体であり，$\Lambda = ||I_\vartheta^{1/2} X - I_\vartheta^{1/2} H_0||^2 - ||I_\vartheta^{1/2} X - I_\vartheta^{1/2} H||^2$ の二項目は 0 になる．よって局所帰無空間 $\sqrt{n}(\Theta_0 - \vartheta)$ は次元 ℓ の線形部分空間に収束し，Lemma から尤度比検定量の漸近分布は自由度 $k-\ell$ のカイ二乗分布となる．