

# Theoretical Statistics まとめノート

Takahiro Nukui

平成 27 年 2 月 16 日

## 目次

<b>1</b>	<b>SOME GENERAL CONCEPTS (2 章)</b>	<b>2</b>
1.1	確率変数の変換 . . . . .	2
1.2	十分統計量 . . . . .	2
1.3	分解定理 . . . . .	2
1.4	最小十分統計量 . . . . .	3
1.5	指数型分布族 . . . . .	5
1.6	完備十分統計量 . . . . .	5
1.7	攪乱母数 (nuisance parameter) と補助統計量 (ancillary statistics) . . . . .	5
<b>2</b>	<b>PURE SIGNIFICANCE TESTS (3 章)</b>	<b>6</b>
2.1	検定の枠組み . . . . .	6
<b>3</b>	<b>SIGNIFICANCE TESTS: SIMPLE NULL HYPOTHESES(4 章)</b>	<b>6</b>
3.1	ネイマン・ピアソンの補題 (Neyman-Pearson lemma) . . . . .	6
<b>4</b>	<b>COMPOSITE NULL HYPOTHESES (5 章)</b>	<b>7</b>
4.1	相似領域 (similar region) とネイマン構造 (Neyman structure) . . . . .	7
<b>5</b>	<b>参考文献</b>	<b>9</b>

# 1 SOME GENERAL CONCEPTS (2章)

## 1.1 確率変数の変換

定理：確率密度関数の変換 (p12)

確率変数  $Y$  が密度関数  $f(Y)$  を持つとき、単調な関数  $h$  で表された  $Z = h(Y)$  の密度関数  $g(Z)$  は、

$$g(z) = f(y) \left| \frac{dy}{dz} \right|$$

証明

$Y$  の密度関数の範囲を  $[a, b]$ ,  $Z$  の密度関数の範囲を  $[c, d]$  とし、 $Z$  の分布関数を  $G(z)$  とする。いま、 $h(x)$  は単調な増加関数とすれば、その逆関数が存在するので、

$$G(z) = P(Z \leq z) = P(h(Y) \leq z) = P(Y \leq h^{-1}(z)) = \int_a^{h^{-1}(z)} f(t) dt = F(y) - F(a)$$

上式を両辺  $z$  で微分して、

$$g(z) = \frac{d}{dz} \{F(y) - F(a)\} = f(y) \frac{dy}{dz}$$

$h$  が単調減少の場合も同様に示せる。(証明終了)

## 1.2 十分統計量

定義：十分統計量 (p19)

統計量  $S$  がパラメータ  $\theta$  に対する十分統計量であるとは、 $S$  を与えたときの条件付き分布が  $\theta$  に依存しないこと。つまり、 $s$  を固定したとき、

$$f_{Y|S}(y|s; \theta)$$

が  $\theta$  によらない。

補足

$S = s$  という情報があれば、 $\theta$  がなんであつても、 $Y = y$  が観測される確率に影響しない。逆に言うと、 $S$  は未知パラメータ  $\theta$  を推定するのに「十分」な情報がある。

## 1.3 分解定理

定理: 分解定理 (p21)

$S$  が  $\theta$  に対して十分統計量であるための必要かつ十分な条件は、ある  $m_1(s, \theta)$  と  $m_2(y)$  という関数が存在して、任意の  $\theta$  に対して、確率密度関数が

$$f_Y(y; \theta) = m_1(s, \theta) m_2(y)$$

と書けること。

証明

(必要性)

$S$  は十分統計量と仮定する。 $S$  は  $Y$  の関数なので、

$$f_Y(y; \theta) = \sum_{S=s(y)} f_{Y|S}(y|s; \theta) = f_{Y|S}(y|s; \theta) = f_S(s; \theta) f_{Y|S}(y|s; \theta)$$

ここで、 $f_{Y|S}(y|s; \theta)$  は  $\theta$  によらないので

$$f_Y(y; \theta) = f_S(s; \theta) f_{Y|S}(y|s)$$

$f_S(s; \theta) = m_1(s, \theta)$ ,  $f_{Y|S}(y|s) = m_2(y)$  と見れば、分解定理の式が成り立つ。

(十分性)

分解定理の式が成り立つと仮定すると、

$$f_{Y|S}(y|s; \theta) = \frac{f_Y(y; \theta)}{f_S(s; \theta)} = \frac{f_Y(y; \theta)}{\sum_{s(z)=s} f_Y(z; \theta)} = \frac{m_1(s, \theta) m_2(y)}{\sum_{s(z)=s} m_1(s, \theta) m_2(z)} = \frac{m_2(y)}{\sum_{s(z)=s} m_2(z)}$$

右辺に  $\theta$  が現れないので、この式は  $\theta$  によらない。よって、 $S$  は十分統計量である。(証明終了)

## 1.4 最小十分統計量

**定義:(弱い意味で) 最小十分統計量**

$T$  を十分統計量とする。

$U = g(T)$  が十分統計量  $\Rightarrow g$  が 1 対 1 写像 (単射)

であれば、 $T$  を (弱い意味で) 最小十分という。

**定義:(強い意味で) 最小十分統計量 (p23)**

$T$  を十分統計量とする。

任意の十分統計量  $S$  に対してある関数  $h$  が存在して  $T = h(S)$  となる

ならば、 $T$  は (強い意味で) 最小十分という。

### 補足

教科書には、強い意味での定義しか書かれていない。弱い意味の最小十分性は、十分性を失うことなく  $T$  の情報をさらに縮約することは不可能であることを意味している。強い意味での最小十分性は、任意の十分統計量の情報を  $T$  に縮約できることを示している。

**定理:弱い意味と強い意味の関係**

強い意味で最小十分ならば、弱い意味でも最小十分である。

また、強い意味の最小十分統計量が存在すれば、弱い意味の最小十分性を満たす  $T$  は強い意味の最小十分性も満たす。

### 証明

(前半の主張の証明)

$T$  が強い意味で最小十分統計量であるとする。 $U = g(T)$  が十分統計量ならば、 $T = h(g(T))$  となる関数  $h$  が存在する。よって、 $g$  は単射である。

(後半の主張の証明)

$T$  が弱い意味で最小十分統計量、 $V$  が強い意味で最小十分統計量であるとする。 $V$  の最小十分性から、ある関数  $h_1$  が存在して、 $V = h_1(T)$  が成り立つ。 $T$  の最小十分性から、 $h$  は 1 対 1 写像だから逆写像が存在し、

$$T = h_1^{-1}(V)$$

が成り立つ。さらに  $V$  の最小十分性から、任意の統計量  $S$  に対して、ある関数  $h_2$  が存在して

$$V = h_2(S)$$

が成り立ち、

$$T = h_1^{-1}(h_2(S))$$

が成り立つ。 $S$  は任意の統計量なので、 $T$  は強い意味でも最小十分である。(証明終了)

同じ統計量を持つ  $y$  は同じ分割領域に所属していると考えれば、十分統計量を定義することは、標本空間の分割を定義することに等しい。最小十分な統計量を定義するような分割の仕方を最小十分統計量の分割と呼ぶとすると、以下の定理が成り立つ。

**定理：強い意味で最小十分統計量を定義する方法 (p24)**

$\theta$  に依存しない関数  $h(z, y)$  で、

$$\mathbf{D}(y) = \{z; \frac{f_Y(z; \theta)}{f_Y(y; \theta)} = h(z, y), \forall \theta \in \Omega\}$$

のように  $Y$  の空間を分割して、統計量を定めたとする<sup>a</sup>。このような分割は、強い意味での最小十分統計量の分割である<sup>b</sup>。

<sup>a</sup> $y_1$  と  $y_2$  が同じ区画に属するならば同じ統計量、異なる区画に属するならば異なる統計量になるように統計量を定義すること。値自体はなんでもいい。

<sup>b</sup> $y \in \mathbf{D}(y_1) \cap \mathbf{D}(y_2)$  ならば、 $\mathbf{D}(y_1) = \mathbf{D}(y_2)$  であることに注意。このため、すべての  $y$  が重複なくどこかの区画に所属する。

**証明**

まず、 $\mathbf{D}(y)$  の分割が十分統計量の分割になっていることを示す。 $y \in \mathbf{D}(y_1)$  とすると、

$$f_Y(y|y \in \mathbf{D}(y_1); \theta) = \frac{f_Y(y; \theta)}{\sum_{y' \in \mathbf{D}(y_1)} f_Y(y'; \theta)} = \frac{h(y, y_1) f_Y(y_1; \theta)}{\sum_{y' \in \mathbf{D}(y_1)} h(y', y_1) f_Y(y_1; \theta)} = \frac{h(y, y_1)}{\sum_{y' \in \mathbf{D}(y_1)} h(y', y_1)}$$

となり、 $\theta$  に依存しないことがわかるので、この分割は十分統計量の定義になっている。

次に、 $\mathbf{D}(y)$  の分割が最小十分統計量の分割になっていることを示す。 $V = v(Y)$  を任意の十分統計量とすると、分解定理によって

$$f_Y(y; \theta) = m_1(v(y), \theta) m_2(y)$$

と書くことができる。したがって、 $v(y) = v(z)$  とすると、

$$\frac{f_Y(y; \theta)}{f_Y(z; \theta)} = \frac{m_2(y)}{m_2(z)}$$

となり、この値は  $\theta$  に依存しない。このことから、

$$v(y) = v(z) \implies y, z \text{ は同じ区画に属する}$$

任意の十分統計量の分割よりも粗い分割であることが示せたので、これは最小十分統計量の分割である。(証明終了)

## 1.5 指数型分布族

定義:指数型分布族 (p27)

$Y_1, Y_2, \dots, Y_n$  が相互に独立で

$$f_{Y_j}(y_j; \theta) = \exp\{a(\theta)b_j(y) + c_j(\theta) + d_j(y)\}$$

つまり、 $Y = (Y_1, \dots, Y_n)$  が

$$f_Y(y; \theta) = \exp\{a(\theta) \sum b_j(y_j) + \sum c_j(\theta) + \sum d_j(y_j)\}$$

と表されるとき、 $f(Y; \theta)$  は指数型分布族をなすという。

### 補足

分解定理より、 $\sum_j b_j(y_j)$  は  $\theta$  に対する十分統計量になっている。

一般には、

$$f_Y(y; \theta) = \exp\left\{\sum_k a_k(\theta)b_k(y) + c(\theta) + d(y)\right\}$$

の形を指数型分布族の定義にしている文献が多そうである。

## 1.6 完備十分統計量

定義:完備十分統計量 (p30)

十分統計量  $S$  が完備であるとは、 $S$  の関数  $h(S)$  で恒等的にその期待値が 0 となるものは、定数 0 に限るとのこと。

すなわち、任意の関数  $h(S)$  に対し、

$$E_S(h(S)) = 0, \forall \theta \implies h(S) \equiv 0$$

が成り立つ<sup>a</sup>ならば、 $S$  は完備である。

$h(S)$  が有界関数のときにのみ上の関係が成り立つならば、特に有界完備 (boundedly complete) とよぶ。

<sup>a</sup> $E_S$  の記号は  $S$  に関して期待値をとるという意味。 $\theta$  を固定して期待値をとるという意味で  $E_\theta$  と表現している文献もあるので注意。

## 1.7 攪乱母数 (nuisance parameter) と補助統計量 (ancillary statistics)

定義:攪乱母数と補助統計量 (p35)

パラメータ  $\theta$  が  $\theta = (\psi, \lambda)$  のように二つの要素に分けて表現できるとする。いま、検定問題において  $\psi$  のみに関心があつて、さしあたって  $\lambda$  に興味のないとき、 $\lambda$  を攪乱母数という。

ここで、 $S$  を  $\theta$  に対する最小十分統計量とし、 $S = (T, C)$  のように二つの要素に分けて表現できるとする。

(a)  $C$  の確率密度関数が  $\lambda$  のみによって決まり、 $\psi$  にはよらず、

(b)  $C = c$  のもとでの  $T$  の確率密度関数が  $\psi$  のみによって決まり、 $\lambda$  にはよらない

ならば、 $C$  を補助統計量といい、 $T$  を  $\psi$  に対して条件付き十分 (conditionally sufficient) という。

### 補足

上記の定義からわかるように、興味のある  $\psi$  について検定するためには、まず  $C$  を観察された  $C = c$  で固定してから、 $T$  の条件付き確率密度関数  $f_T(t|c; \theta)$  を考えればよい。これを Conditionality principle(p38) という。

## 2 PURE SIGNIFICANCE TESTS (3章)

### 2.1 検定の枠組み

**定義:単純仮説と複合仮説 (p64)** —————

帰無仮説  $H_0$  が確率密度関数  $f_Y(y)$  を完全に特定する場合、単純 (simple) 仮説と呼び、そうでなければ複合 (composite) 仮説と呼ぶ。一般に複合仮説は、ある未知パラメータ (攪乱母数) を除いて、確率密度が定まる。

#### 補足

攪乱母数  $\lambda$  に対する十分統計量  $S$  が存在すれば、 $S = s$  を固定することで、 $H_0$  は単純仮説と同一視できるので、嬉しい。このように、帰無仮説が複合仮説のときにはある十分統計量を用いて、攪乱母数によらない分布にもっていくことができる。

## 3 SIGNIFICANCE TESTS: SIMPLE NULL HYPOTHESES(4章)

### 3.1 ネイマン・ピアソンの補題 (Neyman-Pearson lemma)

**定義:最良棄却域 (best critical region)(p91)** —————

帰無仮説を  $H_0$ 、対立仮説を  $H_1$  としたとき、 $P(Y \in w_\alpha; H_0) = \alpha$  を満たす棄却域が、任意の領域  $w_{\alpha'}$  に対して、

$$P(Y \in w_\alpha; H_0) = \alpha \implies P(Y \in w_\alpha; H_1) \geq P(Y \in w_{\alpha'}; H_1)$$

を満たすとき、 $w_\alpha$  を最良棄却域といい、最良棄却域を用いた検定を最強力検定 (most powerful test) という。

**定理:ネイマン・ピアソンの補題 (p92)** —————

二つの仮説

$$\begin{cases} \text{帰無仮説} & H_0 : \theta = \theta_0 \\ \text{対立仮説} & H_1 : \theta = \theta_1 \end{cases}$$

の間で仮説検定を行う際に、

$$w_\alpha = \{y \mid \frac{f_{Y;\theta_0}(y; \theta_0)}{f_{Y;\theta_1}(y; \theta_1)} \leq c_\alpha\}$$

(ただし、定数  $c_\alpha$  は  $P(Y \in w_\alpha; H_0) = \alpha$  となるように設定した値。) がサイズ  $\alpha$  の最良棄却域になる。

#### 証明

証明は教科書 p92。

## 4 COMPOSITE NULL HYPOTHESES (5章)

### 4.1 相似領域 (similar region) とネイマン構造 (Neyman structure)

#### 定義: 相似領域 (p134)

任意の攪乱母数  $\lambda$  に対して、 $Y$  が棄却域  $w_\alpha$  に入る確率が  $\alpha$  で一定ならば、検定がしやすい。このような性質を持つ  $w_\alpha$  を相似領域という。式で書くと

$\forall \lambda \in \Omega_\lambda$  に対して

$$P(Y \in w_\alpha; \phi_0, \lambda) = \alpha$$

#### 定義: ネイマン構造 (p135)

攪乱母数  $\lambda$  に対する十分統計量  $S_\lambda$  があるとき、

$\forall s$  に対して

$$P(Y \in w_\alpha | S_\lambda = s; \phi_0) = \alpha$$

のような性質を持つ棄却域  $w_\alpha$  をネイマン構造という。

#### 定理: 相似領域とネイマン構造の関係 (p135)

$w_\alpha$  がネイマン構造ならば、 $w_\alpha$  は相似領域である。

逆に、 $w_\alpha$  が相似領域で、攪乱母数に対する十分統計量  $S_\lambda$  が完備ならば、 $w_\alpha$  はネイマン構造である。

#### 補足

前半の主張は意味を考えればわかるが、後半の主張の証明は少し技巧的な操作が必要。

#### 証明

( $w_\alpha$  がネイマン構造ならば相似領域であることの証明)

$\forall \lambda \in \Omega_\lambda$  に対して

$$P(Y \in w_\alpha; \phi_0, \lambda) = \int_{S_\lambda} f_{S_\lambda}(s; \phi_0, \lambda) P(Y \in w_\alpha | S_\lambda = s; \phi_0) ds = \int_{S_\lambda} f_{S_\lambda}(s; \phi_0, \lambda) \alpha ds = \alpha$$

( $w_\alpha$  が相似領域で  $S_\lambda$  が完備ならばネイマン構造であることの証明)

指示関数  $I_{w_\alpha}$  を以下のように定義する。

$$I_{w_\alpha}(y) = \begin{cases} 1 & (y \in w_\alpha \text{ のとき}) \\ 0 & (y \notin w_\alpha \text{ のとき}) \end{cases}$$

$w_\alpha$  が相似領域であるとする、 $\forall \lambda \in \Omega_\lambda$  に対して

$$P(Y \in w_\alpha; \phi_0, \lambda) = E_Y\{I_{w_\alpha}(Y); \phi_0, \lambda\} = \alpha$$

つまり、 $\forall \lambda \in \Omega_\lambda$  に対して

$$E_{S_\lambda}[E_Y[I_{w_\alpha}(Y) - \alpha | S_\lambda; \phi_0, \lambda]] = 0 \quad \text{なので}^1,$$

$S$  が有界完備十分統計量であることから、任意の有界関数  $h(s)$  に対して

$$E_{S_\lambda}(h(S)) = 0, \forall \lambda \implies h(s) \equiv 0$$

<sup>1</sup>内側の  $E$  は  $S_\lambda = s$  のもとで、 $Y$  の値を動かして期待値を求めている。外側の  $E$  は、 $S_\lambda$  の値を動かして期待値を求めている。全体では、 $P(Y \in w_\alpha; \phi_0, \lambda) - \alpha$  の値に一致するので 0 になる。

いま、 $h(s) = E_Y[I_{w_\alpha}(Y) - \alpha | S_\lambda = s; \phi_0, \lambda]$  とみれば、 $\forall \lambda$  に対して  $E_{S_\lambda}(h(S)) = 0$  が成り立っていることが示せたので、 $h(s) \equiv 0$  つまり

$$E_Y\{I_{w_\alpha}(Y) | S_\lambda = s; \phi_0, \lambda\} = \alpha, (\forall s)$$

(証明終了)



## 5 参考文献

現代数理統計学 竹村彰通 (創文社)

数理統計学 稲垣宣生 (裳華房)