確率統計基礎講義 B・確率統計特殊講義

2018 年第 1 ターム

目 次

1　仮説検定の基礎理論

1.1　仮説検定とは

例 1.1 (支持率調査)
　共同通信社発表の内閣支持率は 42%
　ある議員「私の体感ではまだ 30%前半だろう」
　支持率 は 35% より大きいと言えるか？
　100 人中 42 人が支持と, 1000 人中 420 人が支持とでは信頼度が異なる.

定義 1.1 (統計的仮説検定問題)
$\mathcal{P} = \{P_\theta;\ \theta \in \Theta\}$
　P_θ : 確率分布
　θ : パラメータ (ベクトル),　Θ : パラメータ空間
　$X \sim P_\theta$: P_θ に従う 確率変数（確率ベクトル）
　$\Theta_0, \Theta_1 \subset \Theta$
　とするとき, X の実現値 x に基づいて

$$H_0 : \theta \in \Theta_0, \quad H_1 : \theta \in \Theta_1$$

のいずれが正しいかを判定する問題を仮説検定問題という.

　例 1.1 の場合, n 人中 X 人が支持すると答えたとすると

$$P_\theta : B(n, \theta) \quad 2\text{ 項分布}, \quad \Theta = [0, 1], \quad \Theta_0 = [0, 0.35], \Theta_1 = (0.35, 1]$$

用語

帰無仮説： データ (実現値) から否定しようとする仮説, あるいは, 誤って正しいと判定した場合の損害が大きい仮説.

対立仮説： 帰無仮説が正しくないと判定される場合に採用される仮説.

棄却： 仮説を正しくないと判定すること.

p 値： 帰無仮説の疑わしさを示す確率（後述）.

有意水準： p 値によって仮説を棄却するときの閾値.

危険率： 第 1 種の過誤確率 (後述) の上限. 有意水準と同じ値.

単純仮説： 仮説が成り立つパラメータがただひとつであるような仮説.

複合仮説： 仮説が成り立つパラメータが複数あるような仮説.

　片側検定　帰無仮説が単純仮説 $H_0 : \theta = \theta_0$ の場合, 対立仮説を

$$H_1 : \theta > \theta_0 \quad \text{あるいは} \quad H_1 : \theta < \theta_0$$

と設定する場合の検定

　片側検定は, 反対側のパラメータの値は起こりえない場合, あるいは, 起こったとしても検出しなくてよい場合に設定する (例：不良品率の検定)

検定の考え方

次の検定問題を考える.

$$X \sim B(n,\theta)\ (2\text{ 項分布}),\quad \text{帰無仮説 } H_0 : \theta = \theta_0 \quad \text{対立仮説 } H_1 : \theta > \theta_0$$

2 種類の過誤

$X \geq k \Rightarrow H_0$ を棄却するとする.

第 1 種の過誤：H_0 が正しいのに, H_0 を棄却する誤り

第 2 種の過誤：H_0 が正しくないのに, H_0 を棄却しない誤り

大数の法則から, n が大きいときは, $\dfrac{X}{n}$ は, θ に近い値と考えられる. したがって, $\dfrac{X}{n} > \theta_0$ で, その差が大きければ大きい程, 仮説は疑わしい. X の実現値 x に対して

$$P_{\theta_0}\{X \geq x\} = \sum_{k=x}^{n} {}_nC_k \theta_0^k (1-\theta_0)^{n-k}$$

を, x の p 値と呼ぶ. 例 1.1 で帰無仮説を $H_0 : \theta = 0.35$ とする.

$n = 100, x = 42$ に対しては

$$\text{p 値} = 0.0593018$$

帰無仮説が正しい場合,17 調査すれば 1 回ぐらいは 42%を超える.

$n = 1000, x = 420$ に対しては

$$\text{p 値} = 1.964932 \times 10^{-6}$$

帰無仮説が正しい場合,50 万回調査すれば 1 回ぐらい 42%を超える.

$$\text{p 値} \leq \alpha \quad \Rightarrow \quad H_0 \text{ 棄却}$$

という検定方法を考えたとき, 疑わしさの基準となる値 α を有意水準とよぶ.

棄却域: 検定方法を定めたとき, 帰無仮説が棄却されるような X の値の集合

例 1.1 で, 実現値 x の p 値を $p(x)$ と書くと

$$p(42) = 0.0593018, \quad p(43) = 0.03891381$$

なので, 有意水準を $\alpha = 0.05$ とするときの棄却域は

$$R = \{x;\ x \geq 43\}$$

となる. このとき, 第 1 種の過誤確率は α 以下となる.

両側検定

$$X \sim B(n,\theta)\ (2\text{ 項分布}),\quad \text{帰無仮説 } H_0 : \theta = \theta_0 \quad \text{対立仮説 } H_1 : \theta \neq \theta_0$$

X/n の値が θ_0 より小さい場合も, その差が大きければ帰無仮説は疑わしいので, 棄却域は, 適当な k, ℓ に対して

$$R = \{x;\ x \leq k \text{ または } x \geq \ell\}$$

が自然. k, ℓ は第 1 種の過誤確率を指定した値 α 以下となるように定めるが, そのような k, ℓ の組み合わせは一意ではない.

$\theta_0 = \dfrac{1}{2}$ の場合, $|x - \dfrac{n}{2}|$ が大きい程帰無仮説は疑わしいので, 実現値 x に対して

$$\text{p 値} = P_{\theta_0}\Big(|X - \frac{n}{2}| \geq |x - \frac{n}{2}|\Big)$$

とするのが自然. $\theta_0 \neq \dfrac{1}{2}$ の場合も

$$\text{p 値} = P_{\theta_0}\Big(|X - n\theta_0| \geq |x - n\theta_0|\Big)$$

がよく用いられている.

1.2　最適化問題

検定問題

$$X \sim P_\theta, \quad \text{帰無仮説 } H_0 : \theta \in \Theta_0, \quad \text{対立仮説 } H_1 : \theta \in \Theta_1, \quad (\Theta_0 \cap \Theta_1 = \emptyset)$$

検出力 帰無仮説が正しくないとき, 帰無仮説を棄却する確率
R を棄却域とする検定に対して

$$\begin{aligned}
&\text{第 1 種の過誤確率} = P_\theta(X \in R) \quad (\theta \in \Theta_0),\\
&\text{第 2 種の過誤確率} = P_\theta(X \in R^c) \quad (\theta \in \Theta_1)\\
&\text{検出力} = 1 - P_\theta(X \in R^c) = P_\theta(X \in R) \quad (\theta \in \Theta_1)
\end{aligned}$$

となる.

$$\beta(\theta; R) = P_\theta(X \in R) \quad (\theta \in \Theta_1)$$

を検出力関数とよぶ.

最適化問題
棄却域 R を小さくすると, 第 1 種の過誤確率は小さくなるかもしれないが, 第 2 種の過誤確率は一般に大きくなる.
指定した α に対して

$$\sup_{\theta \in \Theta_0} P_\theta(X \in R) \leq \alpha$$

の条件下で, 検出力 $\beta(\theta; R)$ を最大とする.

定義 1.2 (一様最強力検定)
有意水準 α の検定 R_0(棄却域) が, 有意水準 α の検定 R に対して

$$P_\theta(X \in R_0) \geq P_\theta(X \in R) \quad \theta \in \Theta_1$$

を満たすとき, R_0 を一様最強力検定と呼ぶ.

定理 1.1 (ネイマンピアソンの基本定理)
検定問題

$$X \sim P_\theta, \quad \text{帰無仮説 } H_0 : \theta = \theta_0, \quad \text{対立仮説 } H_1 : \theta = \theta_1$$

$f(x;\theta)$: P_θ の確率密度関数
　正の数 c に対して

$$\begin{aligned} R_c &= \{x; f(x;\theta_1) \geq c\, f(x;\theta_0)\}, \\ \alpha_c &= \int_{R_c} f(x;\theta_0)dx \end{aligned}$$

とするとき, R_c を棄却域とする検定は, 有意水準 α_c の検定の中で検出力を最大とする.

　対立仮説が単純仮説の場合は, 一様最強力検定を単に最強力検定と呼ぶ.
証明
R を有意水準 α_c の検定とすると

$$\int_R f(x;\theta_0)dx \leq \alpha_c = \int_{R_c} f(x;\theta_0)dx$$

検出力の差は

$$\begin{aligned} &\int_{R_c} f(x;\theta_1)dx - \int_R f(x;\theta_1)dx \\ &\quad \geq \int_{R_c} f(x;\theta_1)dx - \int_R f(x;\theta_1)dx - c\left\{\int_{R_c} f(x;\theta_0)dx - \int_R f(x;\theta_0)dx\right\} \\ &\quad = \int_{R_c} f(x;\theta_1) - cf(x;\theta_0)dx - \int_R f(x;\theta_1) - cf(x;\theta_0)dx \\ &\quad = \int_{R_c \cap R^c} f(x;\theta_1) - cf(x;\theta_0)dx - \int_{R \cap R_c^c} f(x;\theta_1) - cf(x;\theta_0)dx \geq 0 \end{aligned}$$

■

例 1.2
検定問題

$$X \sim N(\theta, 1), \quad \text{帰無仮説 } H_0 : \theta = \theta_0, \quad \text{対立仮説 } H_1 : \theta = \theta_1, \quad (\theta_1 > \theta_0)$$

に対して, 有意水準 α の最強力検定検定を求める.

$$f(x;\theta) = \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2}$$

であるから

$$
\begin{aligned}
& f(x;\theta_1) \geq c\, f(x;\theta_0) \\
& \quad \Leftrightarrow \frac{f(x;\theta_1)}{f(x;\theta_0)} \geq c \\
& \quad \Leftrightarrow \exp\{x(\theta_1-\theta_0) - \frac{1}{2}(\theta_1^2-\theta_0)^2\} \geq c \\
& \quad \Leftrightarrow x-\theta_0 \geq \frac{1}{2}(\theta_1-\theta_0) + \frac{\log c}{\theta_1-\theta_0} = C \text{ とおく}
\end{aligned}
$$

H_0 の下で, $X-\theta_0 \sim N(0,1)$ であるから

$$
\alpha = 1-\Phi(C), \quad \Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-x^2/2}dx
$$

となるように C を定めればよい. この C を $N(0,1)$ の上側 α 点と呼び, z_α と表す. 最強力検定は

$$
x-\theta_0 \geq z_\alpha \Rightarrow \text{ 棄却}
$$

問題 1.1
例 1.2 で求めた最強力検定は
　検定問題

$$
X \sim N(\theta,1), \quad \text{帰無仮説 } H_0: \theta=\theta_0, \quad \text{対立仮説 } H_1: \theta=\theta_1, \quad (\theta_1 < \theta_0)
$$

の第 1 種の過誤確率が α の検定の中で, 検出力が最小であることを示せ.

注.　例 1.2 で求めた検定は, $\theta_1 > \theta_0$ である限り, θ_1 の値に依存していないので,
　片側検定問題

$$
X \sim N(\theta,1), \quad \text{帰無仮説 } H_0: \theta=\theta_0, \quad \text{対立仮説 } H_1: \theta=\theta > \theta_0
$$

の一様最強力検定である.

問題 1.2
両側検定問題

$$
X \sim N(\theta,1), \quad \text{帰無仮説 } H_0: \theta=\theta_0, \quad \text{対立仮説 } H_1: \theta \neq \theta_1
$$

の一様最強力検定は存在しないことを示せ.

1.3　確率化検定

例 1.3
検定問題

$$
X \sim B(10,\theta), \quad \text{帰無仮説 } H_0: \theta=0.5, \quad \text{対立仮説 } H_1: \theta=0.3
$$

の有意水準 5% の検定を考える.

$$R_0 = \{0, 1\}, \quad R_1 = \{0, 2\}, \quad R_2 = \{0, 1, 9.10\}, \quad R_3 = \{0, 1, 2\}$$

による検定の第 1 種の過誤確率は

$$\sum_{k \in R_0} {}_{10}C_k (0.5)^{10} = 0.0107422,$$
$$\sum_{k \in R_1} {}_{10}C_k (0.5)^{10} = 0.0449219,$$
$$\sum_{k \in R_2} {}_{10}C_k (0.5)^{10} = 0.0214844$$
$$\sum_{k \in R_3} {}_{10}C_k (0.5)^{10} = 0.0546875$$

となり, R_0, R_1, R_2 は有意水準 5% の条件を満たす. 検出力は

$$\sum_{k \in R_0} {}_{10}C_k (0.3)^k (1 - 0.3)^{10-k} = 0.149308,$$
$$\sum_{k \in R_1} {}_{10}C_k (0.3)^k (1 - 0.3)^{10-k} = 0.261722,$$
$$\sum_{k \in R_2} {}_{10}C_k (0.3)^k (1 - 0.3)^{10-k} = 0.298617$$

となり, R_2 の検出力が最大. しかし, R_1 は, $x = 2$ のとき H_0 を棄却するのに $x = 1$ では棄却しない. R_2 では, $x = 9$ のとき, H_0 より H_1 を選ぶことになる.

定義 1.3 (確率化検定)
検定問題

$$X \sim P_\theta, \quad \text{帰無仮説 } H_0 : \theta \in \Theta_0, \quad \text{対立仮説 } H_1 : \theta \in \Theta_1, \quad (\Theta_0 \cap \Theta_1 = \emptyset)$$

$\phi(x)$: X の値域 $\mathfrak{X}$ 上の関数で, $0 \le \phi(x) \le 1$
実現値 x に対して, 確率 $\phi(x)$ で H_0 を棄却する検定を, 確率化検定といい, $\phi(x)$ を検定関数と呼ぶ.

U を X と独立に区間 $(0, 1)$ 上の一様分布に従う確率変数とし, (X, U) の実現値 (x, u) によって仮説検定を行うことを考える.

$$(X, U) \sim P_\theta \times U(0, 1), \quad \text{帰無仮説 } H_0 : \theta \in \Theta_0, \quad \text{対立仮説 } H_1 : \theta \in \Theta_1, \quad (\Theta_0 \cap \Theta_1 = \emptyset)$$

棄却域

$$\tilde{R} = \{(x, u);\ u \le \phi(x)\}$$

によって定まる検定は, $\phi(x)$ を検定関数とする確率化検定であり,

$$\mathrm{P}_\theta((X, U) \in \tilde{R}) = \mathrm{P}_\theta(U \le \phi(X)) = \mathrm{E}_\theta[\phi(X)]$$

$R \subset \mathcal{X}$ に対して

$$\phi(x) = \begin{cases} 1 & x \in R \\ 0 & x \notin R \end{cases}$$

と定義すると, 棄却域 R によって定められる (確率化しない) 検定と一致する.

最適化問題
指定した α に対して

$$\sup_{\theta \in \Theta_0} \mathrm{E}[\phi(X)] \leq \alpha$$

の条件下で, 検出力 $\mathrm{E}[\phi(X)]$ $(\theta \in \Theta_1)$ を最大にする.

定理 1.2 (ネイマンピアソンの基本定理：確率化検定)
検定問題

$$X \sim P_\theta, \quad \text{帰無仮説 } H_0 : \theta = \theta_0, \quad \text{対立仮説 } H_1 : \theta = \theta_1$$

$f(x;\theta)$: P_θ の確率密度関数

$\mathrm{P}_{\theta_0}(f(X;\theta_1) > 0) > \alpha$ と仮定する.

正の数 c に対して

$$\phi(x;c,\gamma) = \begin{cases} 1 & f(x;\theta_1) > cf(x;\theta_0) \\ \gamma & f(x;\theta_1) = cf(x;\theta_0) \\ 0 & f(x;\theta_1) < cf(x;\theta_0) \end{cases}$$

とする. c, γ を

$$\mathrm{E}_{\theta_0}[\phi(X;c,\gamma)] = \alpha$$

となるように定めると, 有意水準が α であるような確率化検定の中で検出力が最大となる.

証明
$\psi(x)$ を有意水準 α の任意の確率化検定とすると

$$\int \{\phi(x;c,\gamma) - \psi(x)\} f(x;\theta_0) dx \geq 0$$

だから, 検出力の差は

$$\begin{aligned} &\int \{\phi(x;c,\gamma) - \psi(x)\} f(x;\theta_1) dx \\ &\geq \int \{\phi(x;c,\gamma) - \psi(x)\} f(x;\theta_1) dx - c \int \{\phi(x;c,\gamma) - \psi(x)\} f(x;\theta_0) dx \\ &= \int \{\phi(x;c,\gamma) - \psi(x)\} \{f(x;\theta_1) - c\, f(x;\theta_0)\} dx \geq 0 \end{aligned}$$

■

c, γ の決め方　$\mathrm{P}_{\theta_0}(f(X;\theta_1) > k\, f(X;\theta_0))$ は k の減少関数なので

$$c = \inf\{k;\ \mathrm{P}_{\theta_0}(f(X;\theta_1) > k\, f(X;\theta_0)) \le \alpha\}$$

とすると

$$\delta := \mathrm{P}_{\theta_0}(f(X;\theta_1) > c\, f(X;\theta_0)) \le \alpha,$$
$$c' < c \Rightarrow \mathrm{P}_{\theta_0}(f(X;\theta_1) > c\, f(X;\theta_0)) > \alpha,$$

となる. 等号が成り立つとき, $\gamma = 0$, 成り立たないときは

$$\mathrm{P}_{\theta_0}(f(X;\theta_1) = c\, f(X;\theta_0)) > 0$$

なので

$$\gamma = \frac{\alpha - \delta}{\mathrm{P}_{\theta_0}(f(X;\theta_1) = c\, f(X;\theta_0))}$$

と定めればよい.

例 1.4
検定問題

$$X \sim B(n, \theta), \quad \text{帰無仮説 } H_0 : \theta = \theta_0, \quad \text{対立仮説 } H_1 : \theta = \theta_1\ (\theta_1 > \theta_0)$$

の有意水準 5% の検定を考える.

$$f(x;\theta) = {}_nC_x \theta^x (1-\theta)^{n-x} \quad (x = 0, 1, \ldots, n)$$

であるから

$$\begin{aligned}
&f(x;\theta_1) > c\, f(x;\theta_0) \\
&\Leftrightarrow \frac{f(x;\theta_1)}{f(x;\theta_0)} > c \Leftrightarrow \frac{(\frac{\theta_1}{1-\theta_1})^x}{(\frac{\theta_0}{1-\theta_0})^x}\left(\frac{1-\theta_1}{1-\theta_0}\right)^n > c \\
&\Leftrightarrow x > \frac{\log c}{\log \frac{\theta_1}{1-\theta_1} - \log \frac{\theta_0}{1-\theta_0}} + n \log \frac{1-\theta_0}{1-\theta_1} = m \text{ とおく}
\end{aligned}$$

c の代わりに m を決めれば良いので

$$m = \inf\{k;\ \mathrm{P}_{\theta_0}(X > m) \le \alpha\}$$

とする. このとき, m は整数となり,

$$\mathrm{P}_{\theta_0}(X \ge m+1) \le 0.05,\ \mathrm{P}_{\theta_0}(X \ge m) > 0.05$$

となる.

$$\phi(x) = \begin{cases} 1 & x > m \\ \dfrac{0.05 - \mathrm{P}_{\theta_0}(X > m)}{\mathrm{P}_{\theta_0}(X = m)} & x = m \\ 0 & x < m \end{cases}$$

は, 最強力検定となる. この検定は $\theta_1 > \theta_0$ である限り, θ_1 に依存しないので
　検定問題

$$X \sim B(n, \theta), \quad \text{帰無仮説 } H_0 : \theta = \theta_0, \quad \text{対立仮説 } H_1 : \theta > \theta_0$$

に対する一様最強力検定である.

1.4　不偏検定

検定問題

$$X \sim N(\theta, 1), \quad \text{帰無仮説 } H_0 : \theta = \theta_0, \quad \text{対立仮説 } H_1 : \theta = \theta_1$$

に対する有意水準 α の最強力検定の検定関数は
$\theta_1 > \theta_0$ のとき

$$\phi_+(x) := \begin{cases} 1 & x \geq z_\alpha \\ 0 & x < z_\alpha \end{cases}$$

$\theta_1 < \theta_0$ のとき

$$\phi_-(x) \begin{cases} 1 & x \leq -z_\alpha \\ 0 & x > -z_\alpha \end{cases}$$

検定問題が両側で:

$$X \sim N(\theta, 1), \quad \text{帰無仮説 } H_0 : \theta = \theta_0, \quad \text{対立仮説 } H_1 : \theta \neq \theta_0$$

であるとき, $\theta < \theta_0$ であるとき,

$$\mathrm{E}[\phi_+(X) \,|\, \theta] < \mathrm{E}[\phi_+(X) \,|\, \theta_0]$$

であることがわかるので, ϕ_+ による検定は

H_0 が正しいときにに H_0 を棄却する確率より

H_1 が正しいときに H_0 を棄却する確率の方が小さい

定義 1.4 (不偏検定)
検定問題

$$X \sim P_\theta, \quad \text{帰無仮説 } H_0 : \theta \in \Theta_0, \quad \text{対立仮説 } H_1 : \theta \in \Theta_1 \ (\Theta_0 \cap \Theta_1 = \emptyset)$$

の検定 ϕ(検定関数) が, 任意の $\theta_0 \in \Theta_0, \theta_1 \in \Theta$ に対して

$$\mathrm{E}_{\theta_0}[\phi(X)] \leq \mathrm{E}_{\theta_1}[\phi(X)]$$

を満たすとき, 不偏検定と呼ぶ.

定義 1.5 (一様最強力不偏検定)
検定問題

$$X \sim P_\theta, \quad \text{帰無仮説 } H_0 : \theta \in \Theta_0, \quad \text{対立仮説 } H_1 : \theta \in \Theta_1 \ (\Theta_0 \cap \Theta_1 = \emptyset)$$

に対する不偏検定 の中で, $\theta \in \Theta_1$ について一様に検出力を最大とする検定を一様最強力不偏検定と呼ぶ.

補題 1.1 (ネイマンピアソンの基本定理の拡張)
$f_1(x), f_2(x), \ldots, f_{m+1}(x)$ を $\mathbb{R}$ 上の可積分関数とし, 非負の定数 $k_1, \ldots, k_m$ に対して

$$\phi(x; k_1, \ldots, k_m) = \begin{cases} 1 & f_{m+1}(x) > \sum_{i=1}^m k_i f_i(x) \\ 0 & f_{m+1}(x) < \sum_{i=1}^m k_i f_i(x) \end{cases}$$

と定義する.

$$\alpha_i(k_1, \ldots, k_m) = \int \phi(x; k_1, \ldots, k_m) f_i(x) dx, \quad i = 1, \ldots, m$$

とするとき, $0 \leq \psi(x) \leq 1$ で, かつ

$$\int \psi(x) f_i(x) dx \leq \alpha_i(k_1, \ldots, k_m), \quad i = 1, \ldots, m$$

を満たす任意の関数 $\psi(x)$ に対して

$$\int \phi(x; k_1, \ldots, k_m) f_{m+1}(x) dx \geq \int \psi(x) f_{m+1}(x) dx$$

が成り立つ.
証明

$$\begin{aligned}
&\int \phi(x; k_1, \ldots, k_m) f_{m+1}(x) dx - \int \psi(x) f_{m+1}(x) dx \\
&\geq \int \phi(x; k_1, \ldots, k_m) f_{m+1}(x) dx - \int \psi(x) f_{m+1}(x) dx \\
&\qquad - \sum_{i=1}^m k_i \left\{ \alpha_i(k_1, \ldots, k_m) - \int \psi(x) f_i(x) dx \right\} \\
&= \int \{\phi(x; k_1, \ldots, k_m) - \psi(x)\} \left\{ f_{m+1}(x) - \sum_{i=1}^m k_i f_i(x) \right\} dx \geq 0
\end{aligned}$$

■

例 1.5
検定問題

$$X \sim N(\theta, 1), \quad \text{帰無仮説 } H_0 : \theta = \theta_0, \quad \text{対立仮説 } H_1 : \theta \neq \theta_0$$

に対して $\phi(x)$ が有意水準 α の不偏検定であるための条件は

$$\beta(\theta; \phi) = \mathrm{E}_\theta[\phi(X)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(x) e^{-(x-\theta)^2/2} dx$$

とすると

$$\beta(\theta_0; \phi) = \alpha, \tag{1.1}$$

$$\beta(\theta_0;\phi) \leq \beta(\theta;\phi) \quad ({}^\forall\theta)$$

である. $\beta(\theta;\phi)$ は連続関数なので (check!), $\beta(\theta;\phi)$ は $\theta=\theta_0$ で最小値をとり, したがって

$$\frac{d}{d\theta}\beta(\theta_0;\phi) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\phi(x)(x-\theta_0)e^{-(x-\theta_0)^2/2}dx = 0 \tag{1.2}$$

(1.1), (1.2) の条件下で, $\theta_1 \neq \theta_0$ に対して $\beta(\theta_1;\phi)$ を最大とする検定関数は, 補題 1.1 より

$$\phi(x;a,b) = \begin{cases} 1 & e^{-(x-\theta_1)^2/2} > ae^{-(x-\theta_0)^2/2} + b(x-\theta_0)e^{-(x-\theta_0)^2/2} \\ 0 & e^{-(x-\theta_1)^2/2} < ae^{-(x-\theta_0)^2/2} + b(x-\theta_0)e^{-(x-\theta_0)^2/2} \end{cases}$$

$$\phi(x;a,b) = 1 \Leftrightarrow e^{(2x-\theta_0-\theta_1)(\theta_1-\theta_0)} > a + b(x-\theta_0)$$

$y = e^{(2x-\theta_0-\theta_1)(\theta_1-\theta_0)}$ は下に凸の曲線であるから, 直線 $y = a+b(x-\theta_0)$ との交点は高々2個であり, 交点の個数が 1 個以下のときは

$$\mathrm{E}_{\theta_0}[\phi(X;a,b)] = 1$$

となるので, $\alpha < 1$ ならば, 不等式は, 適当な $A < B$ をを用いて

$$x - \theta_0 < A \quad \text{または} \quad B < x - \theta_0 \tag{1.3}$$

と表される. (a,b) と (A,B) は 1 対 1 対応であるから (1.3) を棄却域とする検定で (1.1), (1.2) を満たすものを求めればよい.

$$A = -z_{\alpha/2}, \quad B = z_{\alpha/2}$$

ととると, (1.1), (1.2) が成り立ち, θ_1 に依存しないので

$$|X-\theta_0| > z_{\alpha/2} \quad \Rightarrow \quad H_0 \text{ を棄却}$$

は, 一様最強力不偏検定である.

1.5　尤度比検定

定義 1.6 (尤度比検定)

$X \sim P_\theta$ とし, $f(x;\theta)$ を P_θ の確率密度関数とする. 実現値 x に対して

$$L(\theta;x) = f(x;\theta)$$

を, 尤度関数という.

検定問題

$$X \sim P_\theta, \quad \text{帰無仮説 } H_0 : \theta \in \Theta_0, \quad \text{対立仮説 } H_1 : \theta \in \Theta_1, \quad (\Theta_0 \cap \Theta_1 = \emptyset)$$

に対して,

$$\frac{\max_{\theta\in\Theta_0} L(\theta;x)}{\max_{\theta\in\Theta_1} L(\theta;x)} < c \quad \Rightarrow \quad H_0 \text{ を棄却}$$

あるいは,

$$\frac{\max_{\theta\in\Theta_0} L(\theta;x)}{\max_{\theta\in\Theta_0\cup\Theta_1} L(\theta;x)} < c \quad \Rightarrow \quad H_0 \text{ を棄却}$$

を尤度比検定と呼ぶ.

例 1.6
検定問題

$$\begin{aligned}
&X_1,\ldots,X_n \overset{i.i.d.}{\sim} N(\mu,\sigma^2), \quad H_0:\mu=\mu_0 \quad \text{vs} \quad H_1:\mu\neq\mu_0 \\
&(\Theta=\{(\mu,\sigma^2);\mu\in\mathbb{R},\sigma^2>0\}, \quad \Theta_0=\{(\mu_0,\sigma^2);\mu\in\sigma^2>0\},)
\end{aligned}$$

$$L(\mu,\sigma^2;x_1,\ldots,x_n)=(2\pi\sigma^2)^{-n/2}\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu)^2\}$$

$$\begin{aligned}
&\hat\sigma_0^2 := \operatorname*{argmax}_{\sigma^2} L(\mu_0,\sigma^2;x_1,\ldots,x_n)=\frac{1}{n}\sum_{i=1}^n(x_i-\mu_0)^2, \\
&L(\mu_0,\hat\sigma_0^2;x_1,\ldots,x_n)=(2\pi\hat\sigma_0^2)^{-n/2}e^{-n/2} \\
&\operatorname*{argmax}_{\mu} L(\mu,\sigma^2;x_1,\ldots,x_n)=\frac{1}{n}\sum_{i=1}^n x_i =: \bar x, \\
&\hat\sigma^2 := \operatorname*{argmax}_{\sigma^2} L(\bar x,\sigma^2;x_1,\ldots,x_n)=\frac{1}{n}\sum_{i=1}^n(x_i-\bar x)^2, \\
&L(\bar x,\hat\sigma^2;x_1,\ldots,x_n)=(2\pi\hat\sigma^2)^{-n/2}e^{-n/2}
\end{aligned}$$

となる.

$$\hat\sigma_0^2=\frac{1}{n}\sum_{i=1}^n(x_i-\bar x+\bar x-\mu_0)^2=\hat\sigma^2+(\bar x-\mu_0)^2$$

より

$$\frac{\max_{\theta\in\Theta_0} L(\theta;x)}{\max_{\theta\in\Theta} L(\theta;x)} < c \Leftrightarrow \frac{(\bar x-\mu_0)^2}{\hat\sigma^2} > C$$

帰無仮説 H_0 の下で

$$\frac{(\bar X-\mu_0)^2}{\hat\sigma^2/(n-1)}=\frac{n(\bar X-\mu_0)^2}{\sum_{i=1}^n(X_i-\bar X)^2/(n-1)} \sim F_{1,n-1}$$

なので, $C=(n-1)f_{1,n-1}(\alpha)$ とすれば良い, ただし, $f_{1,n-1}(\alpha)$ は, 自由度 $1,n-1$ の F 分布の上側 α 点.

2　点推定の基礎理論

2.1　一様最小分散不偏推定量

定義 2.1 (点推定)

$\mathcal{P} = \{P_\theta;\ \theta \in \Theta\}$

P_θ : 確率分布

θ : パラメータ (ベクトル),　$\Theta \subset \mathbb{R}$: パラメータ空間

$X \sim P_\theta$: P_θ に従う 確率変数（確率ベクトル）

$\gamma : \Theta \to \mathbb{R}^p$

とするとき, X の実現値 x の関数 $g(x)$ の値によって $\gamma = \gamma(\theta)$ を推定すること.

$g(x)$ を推定値, $g(X)$ を推定量という.

<u>推定量の良さの基準</u>$(p = 1)$

平均 2 乗誤差 $\mathrm{E}_\theta[\{g(X) - \gamma\}^2]$

平均絶対誤差 $\mathrm{E}_\theta[|g(X) - \gamma|]$

集中確率 $\mathrm{P}_\theta(|g(X) - \gamma| < \epsilon)$　(ϵ は誤差の許容範囲)

※. 平均 2 乗誤差や平均絶対誤差を θ に関して一様に最小とする推定量は存在しない.

例. $g_1(x) \equiv \gamma_1, g_2(x) \equiv \gamma_2$

$$\gamma(\theta) = \gamma_1 \Rightarrow r(g_1, \theta) = 0 < r(g_2, \theta)$$
$$\gamma(\theta) = \gamma_2 \Rightarrow r(g_2, \theta) = 0 < r(g_1, \theta)$$

ただし, $r(g, \theta) = \mathrm{E}_\theta[\{g(X) - \gamma\}^2]$ または $\mathrm{E}_\theta[|g(X) - \gamma|]$.

定義 2.2 (不偏性)

$$^\forall\theta\ \mathrm{E}_\theta[g(X)] = \gamma$$

を満たす推定量 $g(X)$ を γ の不偏推定量と呼ぶ.

定義 2.3 (一様最小分散不偏推定量 (UMVUE))

不偏推定量の中で, θ に関して一様に分散を最小にする推定量.

UMVU : Uniformly Minimum Variance Unbiased Estimator

定理 2.1 (クラーメルーラオの不等式)

X の確率密度関数を $f(x, \theta)$ とする. $\gamma(\theta)$ は微分可能な関数で

(A) 各 θ_0 に対して, θ_0 の近傍 U と関数 $M(x)$ が存在して, $\theta \in U$ ならば

$$\left\| \frac{\partial f(x, \theta)}{\partial \theta} \right\| \leq M(x),\ \text{かつ},\ \mathrm{E}_{\theta_0}\left[\left\{ \frac{M(X)}{f(X, \theta_0)} \right\}^2 \right] < \infty$$

と仮定する．このとき, $\gamma(\theta)$ の任意の不偏推定量 $g(X)$ に対して

$$\mathrm{Var}\{g(X)\} \geq \dot{\gamma}(\theta)' J(\theta)^{-1} \dot{\gamma}(\theta) \tag{2.4}$$

が成り立つ．ただし, $\dot{\gamma}(\theta) = \partial\gamma(\theta)/\partial\theta$,

$$J(\theta) = \mathrm{E}\left[\left\{\frac{\partial}{\partial\theta}\log f(X,\theta)\right\}\left\{\frac{\partial}{\partial\theta}\log f(X,\theta)\right\}'\right] \tag{2.5}$$

であり, $J(\theta)$ は正則であると仮定する．

証明
X が連続型の場合に証明する．離散型の場合は, 積分を和に置き換えればよい．$\mathrm{Var}\{g(X)\} = \infty$ のとき, 示したい不等式は自明なので, $\mathrm{Var}\{g(X)\} < \infty$ とする．

$f(x,\theta)$ は確率密度関数であるので $\displaystyle\int f(x,\theta)dx = 1$ を満たす．この両辺を θ に関して微分する．条件 (A) より, 微分と積分の順序変更ができて

$$\int \frac{\partial f(x,\theta)}{\partial\theta}dx = \mathrm{E}\left[\frac{\partial \log f(X,\theta)}{\partial\theta}\right] = 0. \tag{2.6}$$

次に, $g(x)$ が $\gamma(\theta)$ の不偏推定量であるから

$$\gamma(\theta) = \int g(x)f(x,\theta)dx$$

を満たしている．この両辺を θ に関して微分することを考えるが, シュワルツの不等式より

$$\left\{\int g(x)M(x)dx\right\}^2 \leq \int g(x)^2 f(x,\theta_0)dx \int \left\{\frac{M(x)}{f(x,\theta_0)}\right\}^2 f(x,\theta_0)dx < \infty.$$

よって, 微分と積分の順序を変更し, (2.6) を用いると

$$\dot{\gamma}(\theta) = \int g(x)\frac{\partial \log f(x,\theta)}{\partial\theta}f(x,\theta)dx = \mathrm{Cov}\left\{g(X), \frac{\partial \log f(X,\theta)}{\partial\theta}\right\}$$

を得る．したがって, 任意のベクトル $\boldsymbol{a}$ に対して

$$\boldsymbol{a}'\dot{\gamma}(\theta) = \mathrm{Cov}\left\{g(X), \boldsymbol{a}'\frac{\partial \log f(X,\theta)}{\partial\theta}\right\}$$

が成り立ち, さらに (共分散を内積として) シュワルツの不等式を用いると

$$(\boldsymbol{a}'\dot{\gamma}(\theta))^2 \leq \mathrm{Var}\{g(X)\}\mathrm{Var}\left\{\boldsymbol{a}'\frac{\partial \log f(X,\theta)}{\partial\theta}\right\} = \mathrm{Var}\{g(X)\}\boldsymbol{a}'J(\theta)\boldsymbol{a}$$

を得る．$\boldsymbol{a} = J(\theta)^{-1}\dot{\gamma}(\theta)$ ととることにより不等式 (2.4) が得られる． ■

定義 2.4 (有効推定量)
クラメール－ラオの不等式の下限を達成する不偏推定量を有効推定量と呼ぶ.

例 2.1
$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$

$$\theta = (\mu, \sigma^2)',$$

$$f(x_1, \ldots, x_n; \theta) = (2\pi\sigma^2)^{-n/2} \exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\},$$

$$\log f(x_1, \ldots, x_n; \theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\frac{\partial \log f(X_1, \ldots, X_n; \theta)}{\partial \theta} = \Big[\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu), -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(X_i - \mu)^2\Big)\Big]'$$

$$= \Big[\frac{1}{\sigma}\sum_{i=1}^{n} Z_i, -\frac{1}{2\sigma^2}\Big(n - \sum_{i=1}^{n} Z_i^2\Big)\Big]', \quad Z_i = \frac{1}{\sigma}(X_i - \mu) \sim N(0, 1)$$

$$J(\theta) = \begin{pmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{pmatrix}$$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \mathrm{E}[\bar{X}] = \mu, \quad \mathrm{Var}[\bar{X}] = \frac{\sigma^2}{n}$$

$$S = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \quad \mathrm{E}[S] = \sigma^2, \quad \mathrm{Var}[S] = \frac{2\sigma^4}{(n-1)}$$

$\bar{X}$ は μ の有効推定量であるが, S はσ^2 の有効推定量ではない.

問題 2.1

$$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(0, \theta) \quad (\theta > 0)$$

とする. $S_0 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ は θ の有効推定量であることを示せ.

2.2　十分統計量

定義 2.5 (十分統計量)
$X = (X_1, \ldots, X_n)' \sim P_\theta,\ \theta \in \Theta$
　$T = T(X)$ は, $T = t$ を与えたときの X の条件付き分布が θ に無関係であるとき, θ の十分統計量であるとする.

例 2.2 (2 項分布)
$X_1, \ldots, X_n \overset{i.i.d.}{\sim} B(1, p), \quad T = \sum_{i=1}^{n} X_i$

　$x_i = 0$ or 1 $(i = 1, \ldots, n), \quad \sum_{i=1}^{n} x_i = t$ のとき

$$\mathrm{P}(X_1 = x_1, \ldots, X_n = x_n \,|\, T = t) = \frac{\mathrm{P}(X_1 = x_1, \ldots, X_n = x_n, T = t)}{\mathrm{P}(T = t)}$$

$$= \frac{p^t(1-p)^{n-t}}{{}_nC_t p^t(1-p)^{n-t}} = \frac{1}{{}_nC_t}$$

定理 2.2 (ラオーブラックウェルの定理)
$X = (X_1, \ldots, X_n)' \sim P_\theta,\ \theta \in \Theta$
$T = T(X)$: 十分統計量, $g(X)$: $\gamma(\theta)$ の不偏推定量
$g^*(T) = \mathrm{E}\{g(X) \,|\, T\}$ と定義すると,

$$\mathrm{E}[g^*(T)] = \gamma(\theta) \text{ かつ } \mathrm{Var}\{g^*(T)\} \leq \mathrm{Var}\{g(X)\}$$

証明

$$\mathrm{E}\{g^*(T)\} = \mathrm{E}[\mathrm{E}[g(X) \,|\, T]] = \mathrm{E}[g(X)] = \gamma(\theta)$$

$$\begin{aligned}
&\mathrm{E}[\{g(X) - g^*(T)\}\{g^*(T) - \gamma(\theta)\}] \\
&\quad = \mathrm{E}[\mathrm{E}[\{g(X) - g^*(T)\}\{g^*(T) - \gamma(\theta)\} \,|\, T]] \\
&\mathrm{E}[\{g(X) - g^*(T)\}\{g^*(T) - \gamma(\theta)\} \,|\, T] \\
&\quad = \{\mathrm{E}[g(X) \,|\, T] - g^*(T)\}\{g^*(T) - \gamma(\theta)\} = 0
\end{aligned}$$

より

$$\begin{aligned}
&\mathrm{Var}\{g(X)\} = \mathrm{E}[\{g(X) - \gamma(\theta)\}^2] \\
&\quad = \mathrm{E}[\{g(X) - g^*(T) + g^*(T) - \gamma(\theta)\}^2] \\
&\quad = \mathrm{E}[\{g(X) - g^*(T)\}^2] + \mathrm{E}[\{g^*(T) - \gamma(\theta)\}^2] \\
&\qquad\quad + 2\mathrm{E}[\{g(X) - g^*(T)\}\{g^*(T) - \gamma(\theta)\}] \\
&\quad = \mathrm{E}[\{g(X) - g^*(T)\}^2] + \mathrm{Var}\{g^*(T)\}
\end{aligned}$$

■

例 2.3 (2 項分布 (続き))
$T = \sum_{i=1}^{n} X_i$
$\hat{\theta}_1 = X_1$ は θ の不偏推定量

$$\begin{aligned}
&\mathrm{P}(X_1 = 1 \,|\, T = t) = \sum_{x_1 = 1, \sum_{i=1}^n x_i = t} \frac{1}{{}_nC_t} = \frac{{}_{n-1}C_{t-1}}{{}_nC_t} = \frac{t}{n} \\
&\hat{\theta}_* = \mathrm{E}[\hat{\theta}_1 \,|\, T] = \frac{T}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i \\
&\mathrm{Var}\hat{\theta}_1 = \theta(1-\theta), \quad \mathrm{Var}\hat{\theta}_* = \frac{1}{n}\theta(1-\theta)
\end{aligned}$$

$\hat{\theta} = X_2, \dfrac{X_1 + X_2}{2}$ から $\hat{\theta}_*$ を計算しても同じ結果.

定義 2.6 (完備十分統計量)
$X = (X_1, \ldots, X_n)' \sim P_\theta,\ \theta \in \Theta$

十分統計量 T が任意のボレル可測関数 g に対して

$$({}^\forall\theta \in \Theta\ \mathrm{E}_\theta[g(T)] = 0) \quad \Rightarrow \quad (\text{確率 } 1 \text{ で } g(T(X)) = 0)$$

が成り立つとき, T は完備であるという.

例 2.4 (2 項分布 (続き))
$T = \sum_{i=1}^{n} X_i \sim B(n, \theta)$
$\mathrm{E}_\theta[g(T)] = 0$ とすると

$$\sum_{k=0}^{n} \binom{n}{k} g(k)\theta^k(1-\theta)^{n-k} = 0$$

$$\theta \neq 1 \Rightarrow \sum_{k=0}^{n} \binom{n}{k} g(k)\Big(\frac{\theta}{1-\theta}\Big)^k = 0$$

t に関する高々 n 次の方程式

$$\sum_{k=0}^{n} \binom{n}{k} g(k) t^k = 0$$

が $n+1$ 個以上の解を持つので,

$$\binom{n}{k} g(k) = 0 \quad k = 0, 1, \ldots, n$$

したがって, T は完備十分統計量である.

定理 2.3
$T(X)$: 完備十分統計量
$g(T)$: $\gamma(\theta)$ の不偏推定量
$\Rightarrow g(T)$ は $\gamma(\theta)$ の一様最小分散不偏推定量

証明
$h(X)$ を 不偏推定量とすると任意の θ に対して

$$\mathrm{E}[\mathrm{E}[h(X) \,|\, T] - g(T)] = \gamma(\theta) - \gamma(\theta) = 0$$

であるから $g(T) = \mathrm{E}[h(X) \,|\, T]$. したがって

$$\mathrm{Var}\{g(T)\} = \mathrm{Var}\{\mathrm{E}[h(X) \,|\, T]\} \leq \mathrm{Var}\{h(X)\}$$

■

例 2.5
$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$
$T = (T_1, T_2) = (\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2)$ は完備十分統計量

$$S = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{n}{n-1}\left\{\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \Big(\frac{1}{n}\sum_{i=1}^{n} X_i\Big)^2\right\}$$
$$= \frac{1}{n-1}T_2 - \frac{1}{n(n-1)}T_1^2$$

定理 2.4 (分解定理)
$f(x_1,\ldots,x_n;\theta)$: $X=(X_1,\ldots,X_n)'$ の pdf
　$T(X)$ が θ の十分統計量であるための必要十分条件は, f が非負値関数 g,h を用いて

$$f(x_1,\ldots,x_n;\theta)=g(x_1,\ldots,x_n)h\{T(x);\theta\} \tag{2.7}$$

と表されることである。

証明
(離散型の場合) T が十分統計量であるとする. 一般に

$$\begin{aligned}f(x_1,\ldots,x_n;\theta)&=\mathrm{P}_\theta(X_1=x_1,\ldots,X_n=x_n)\\&=\mathrm{P}_\theta(X_1=x_1,\ldots,X_n=x_n|T(x)=t)\cdot\mathrm{P}_\theta(T(X)=t)\end{aligned}$$

と表せる. ここで

$$g(x_1,\ldots,x_n)=\mathrm{P}_\theta(x=x|T(x)=t),\quad h(t;\theta)=\mathrm{P}_\theta(T(X)=t)$$

と定義すると, (2.7) が成り立つ.
　逆に (2.7) を仮定すると

$$\mathrm{P}_\theta(T(X)=t)=\sum_{\boldsymbol{x};T(\boldsymbol{x})=t}g(x_1,\ldots,x_n)h(T(x);\theta)=h(t;\theta)\sum_{\boldsymbol{x};T(\boldsymbol{x})=t}g(x_1,\ldots,x_n)$$

より, $T(T)=t$ が与えられたときの X の条件付き確率密度関数は

$$\frac{\mathrm{P}_\theta(X_1=x_1,\ldots,X_n=x_n,T(X)=t)}{\mathrm{P}_\theta(T(X)=t)}=\frac{g(x_1,\ldots,x_n)}{\sum_{\boldsymbol{x};T(\boldsymbol{x})=t}g(x_1,\ldots,x_n)},\quad(T(\boldsymbol{x})=t)$$

となり, θ に依存しない.
(連続型の場合) 逆変換が C^1 クラスであるような 1 対 1 変換 $y_1=y_1(x_1,\ldots,x_n),\ldots,y_n=y_n(x_1,\ldots,x_n)$ で, 次を満たすものが存在するとして, 証明する.

$$y_n(x_1,\ldots,x_n)=T(x),\quad J(\boldsymbol{y})=\begin{vmatrix}\dfrac{\partial x_1}{\partial y_1}&\cdots&\dfrac{\partial x_1}{\partial y_n}\\\vdots&\ddots&\vdots\\\dfrac{\partial x_n}{\partial y_1}&\cdots&\dfrac{\partial x_n}{\partial y_n}\end{vmatrix}\neq0,\quad{}^\forall\boldsymbol{y}\in\boldsymbol{y}(\mathcal{X}),$$

ただし, $\mathcal{X}=\{x\,;\,f(x;\theta)>0\}$, $\boldsymbol{y}(\mathcal{X})$ は変換 $\boldsymbol{y}(x)=(y_1(x),\ldots,y_n(x))'$ による $\mathcal{X}$ の像である.
　T が十分統計量であると仮定する. 逆変換を $x(\boldsymbol{y})=(x_1(\boldsymbol{y}),\ldots,x_n(\boldsymbol{y}))'$ と表すと, $\boldsymbol{Y}=\boldsymbol{y}(x)$ の同時確率密度関数は

$$f_{\boldsymbol{Y}}(\boldsymbol{y};\theta)=f(x_1(\boldsymbol{y}),\ldots,x_n(\boldsymbol{y});\theta)|J(\boldsymbol{y})|$$

によって与えられる. $T=Y_n=t$ が与えられたときの, $Y_1,\ldots,Y_{n-1}$ の条件付き確率密度関数を $g_{\boldsymbol{Y}}(y_1,\ldots,y_{n-1}|\,t)$, Y_n の周辺確率密度関数を $f_{\boldsymbol{Y}}(y_n;\theta)$ とおくと

$$f_{\boldsymbol{Y}}(\boldsymbol{y};\theta)=f_{\boldsymbol{Y}}(y_n;\theta)g_{\boldsymbol{Y}}(y_1,\ldots,y_{n-1}|\,y_n)$$

となるから

$$f(x_1(\boldsymbol{y}),\ldots,x_n(\boldsymbol{y});\theta)=f_{\boldsymbol{Y}}(y_n;\theta)g_{\boldsymbol{Y}}(y_1,\ldots,y_{n-1}|\,y_n)\frac{1}{|J(\boldsymbol{y})|}.$$

したがって

$$f(x;\theta)=f_{\boldsymbol{Y}}(T(x);\theta)g_{\boldsymbol{Y}}(y_1(x),\ldots,y_{n-1}(x)|\,t(x))\frac{1}{|J(\boldsymbol{y}(x))|}$$

となり, (2.7) が成り立つ.

逆に, (2.7) が成り立つとすると $\boldsymbol{Y}$ の同時確率密度関数は

$$f_{\boldsymbol{Y}}(\boldsymbol{y};\theta)=g(x(\boldsymbol{y}))h(y_n;\theta)|J(\boldsymbol{y})|$$

となり, $y_n=t$ が与えられたときの, $Y_1,\ldots,Y_{n-1}$ の条件付き確率密度関数は

$$\frac{f_{\boldsymbol{Y}}(\boldsymbol{y};\theta)}{\iint_{\mathcal{Y}_{n-1}}f_{\boldsymbol{Y}}(\boldsymbol{y};\theta)dy_1\cdots dy_{n-1}}=\frac{g(x(\boldsymbol{y}))|J(\boldsymbol{y})|}{\iint_{\mathcal{Y}_{n-1}}g(x(\boldsymbol{y}))|J(\boldsymbol{y})|dy_1\cdots dy_{n-1}}$$

と表される. ここで, $\mathcal{Y}_{n-1}=\{(y_1,\ldots,y_{n-1})';\ {}^{\exists}y_n\ s.t.\ (y_1,\ldots,y_{n-1},y_n)'\in\mathcal{Y}\}$ である. 変換 $\boldsymbol{y}(x)$ は 1 対 1 変換であるので, x の $T=t$ を与えたときの条件付き分布は, $y_n=t$ を与えたときの $Y_1,\ldots,Y_{n-1}$ 条件付き分布によって定まり, θ に依存しない. ■

定義 2.7 (指数型分布族)

$\mathcal{P}=\{P_\theta;\ \theta\in\Theta\}$ とし, $f(x;\theta)$ を P_θ の確率密度関数とする.

$$f(x;\theta)=\exp\Big\{\sum_{i=1}^{p}\eta_i(\theta)T_i(x)+D(\theta)\Big\}h(x)$$

であるとき, $\mathcal{P}$ を指数型分布族という. ただし,

$$\begin{aligned}
&x\in\mathbb{R}^q,\\
&\theta=(\theta_1,\ldots,\theta_p)'\in\Theta\subset\mathbb{R}^q,\ \text{開集合}\\
&T_1,\ldots,T_p,h:\mathbb{R}^q\ \text{上で定義された可測関数},\ h(x)\geq 0\\
&D:\Theta\ \text{上で定義された関数}
\end{aligned}$$

注 .

- $\eta=(\eta_1(\theta),\ldots,\eta_p(\theta))'$ を natural parameter と呼ぶ.
- $D(\theta)$ は η の値のみで決まる

$$e^{D(\theta)}=\Bigg\{\int_{\mathbb{R}^q}\exp\Big\{\sum_{i=1}^{p}\eta_i(\theta)T_i(x)\Big\}h(x)dx\Bigg\}^{-1}$$

例 2.6

ガンマ分布

$$f(x;\theta)=\begin{cases}\dfrac{1}{\Gamma[r]}\lambda^r e^{-\theta x}x^{r-1} & x>0\\ 0 & x\leq 0\end{cases}$$

ポアソン分布

$$f(x;\theta) = \begin{cases} e^{-\theta}\dfrac{\theta^x}{x!} & x = 0, 1, 2, \ldots \\ 0 & \text{その他} \end{cases}$$

正規分布

$$f(x;\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \theta = (\mu, \sigma^2)$$

定理 2.5
$\mathcal{P}$: 定義 2.7 で与えられた確率密度関数を持つ指数型分布族
　$X \sim P_\theta \in \mathcal{P}$
　　　$\Rightarrow T(X) = (T_1(X), \ldots, T_p(X))'$ は十分統計量
　さらに, $p \leq q$ で, かつ $\eta(\Theta)$ が $\mathbb{R}^p$ の開集合を含む
　　　$\Rightarrow T(X)$ は θ の完備十分統計量.

証明 $p = 1$ の場合

$$X \sim \exp\{\eta T(x) + D(\eta(\theta))\}h(x)$$
$$\mathrm{E}_\eta[g\{T(X)\}] = 0 \Rightarrow \int g^+(T(x))e^{\eta T(x)}h(x)dx = \int g^-(T(x))e^{\eta T(x)}h(x)dx = c_\eta$$
$$\mathrm{P}_\theta(g((X)) = 0) < 1 \Rightarrow c_\eta > 0$$
$$\eta_0 : \eta(\Theta) \text{ の内点}, \quad h^\pm(x) := \frac{1}{c_{\eta_0}} g^\pm(T(x))e^{\eta_0 T(x)}h(x)$$
$$\eta = \eta_0 + t \in \eta(\Theta)$$
$$\int g^\pm(T(x))e^{\eta T(x)}h(x)dx = c_{\eta_0}\int e^{tT(x)}h^\pm(x)dx$$
$$\int e^{tT(x)}h^+(x)dx = \int e^{tT(x)}h^-(x)dx$$

$X \sim h^+(x)$ のときの $T(X)$ の分布と $X \sim h^-(x)$ のときの $T(X)$ の分布は同じ

$$A^+ = \{x \,|\, g(T(x)) \geq 0\}, A^- = \{x \,|\, g(T(x)) < 0\}$$
$$\mathrm{P}(T(X) \in A^+) = \int_{A^+} h^-(x)dx = 0, \mathrm{P}(T(X) \in A^-) = \int_{A^-} h^+(x)dx = 0$$

$\Rightarrow \mathrm{P}(T(X) \in \mathbb{R}) = 0$ となり矛盾

■

3 区間推定

定義 3.1 (信頼区間)
$X \sim P_\theta, \quad \theta \in \Theta$
　$\gamma : \Theta \to \mathbb{R}$

$0 < \alpha < 1$,
$T_1(X), T_2(X)$: 統計量

$$^\forall\theta,\ P_\theta(T_1(X) \leq \gamma(\theta) \leq T_2(X)) \geq 1 - \alpha$$

を満たすとき, 区間 $[T_1(X), T_2(X)]$ を水準 $1-\alpha$ の $\gamma(\theta)$ の信頼区間という.

定義 3.2 (信頼領域)
$X \sim P_\theta, \quad \theta \in \Theta$
$\gamma : \Theta \to \mathbb{R}^q$
$0 < \alpha < 1$,
$R : X \to R_X \in \wp(\gamma(\Theta))$

$$^\forall\theta,\ P_\theta(\gamma(\theta) \in R_X) \geq 1 - \alpha$$

を満たすとき, 区間 R_X を水準 $1-\alpha$ の $\gamma(\theta)$ の信頼領域という.

信頼区間の構成 (その 1) $g(X)$: $\gamma(\theta)$ の推定量

$$\frac{g(X) - \gamma(X)}{S(X)}$$

の分布が θ に依存しないような統計量 $S(X)$ が存在するとき,

$$\mathrm{P}\left(\left|\frac{g(X) - \gamma(X)}{S(X)}\right| \leq d\right) \geq 1 - \alpha$$

となる d を用いて

$$T_1(X) = g(X) - d\,S(X), T_2(X) = g(X) + d\,S(X)$$

と定めれば, $[T_1(X), T_2(X)]$ が 水準 $1-\alpha$ の信頼区間となる.

例 3.1 (母平均の信頼区間)
$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$
μ の信頼区間を構成する.

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_i : \mu \text{ の不偏推定量 },$$

$$\bar{X} - \mu \sim N(0, \frac{1}{n}\sigma^2)$$

<u>σ^2 が既知の場合</u>

$$Z = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0, 1)$$

$$1 - \alpha = \mathrm{P}(|Z| \leq z_{\alpha/2}) = \mathrm{P}\left(|\bar{X} - \mu| \geq z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}\right)$$

$$T_1 = \bar{X} - z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}, \quad T_2 = \bar{X} + z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}$$

σ^2 が未知の場合

$$S = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 : \sigma^2 \text{ の不偏推定量}$$
$$T = \frac{\sqrt{n}(\bar{X}-\mu)}{\sqrt{S}} \sim t_{n-1}$$
$$1-\alpha = \mathrm{P}(|T| \le t_{n-1}(\alpha/2)) = \mathrm{P}\left(|\bar{X}-\mu| \ge t_{n-1}(\alpha/2)\sqrt{\frac{S}{n}}\right)$$
$$T_1 = \bar{X} - t_{n-1}(\alpha/2)\sqrt{\frac{S}{n}}, \quad T_2 = \bar{X} + t_{n-1}(\alpha/2)\sqrt{\frac{S}{n}}$$

信頼領域の構成　S_X : 水準 α の $\gamma(\theta)$ の信頼領域
　検定問題

$$H_0 : \gamma(\theta) = \gamma_0 \text{ v.s. } H_1 : \gamma(\theta) \neq \gamma_0$$

$$\phi(x) = \begin{cases} 1 & \gamma_0 \notin S_x \\ 0 & \gamma_0 \in S_x \end{cases}$$

は, 有意水準 α の検定となる. $\gamma(\theta) = \gamma_0$ のとき

$$\mathrm{E}_\theta[\phi(X)] = \mathrm{P}_\theta(\gamma(\theta) \notin S_X) = 1 - \mathrm{P}_\theta(\gamma(\theta) \in S_X) \le \alpha$$

　逆に

$$H_0 : \gamma(\theta) = \gamma \text{ v.s. } H_1 : \gamma(\theta) \neq \gamma$$

の検定関数 $\phi(x;\gamma)$ が与えられるとき,

$$S_x = \{\gamma;\ \phi(x;\gamma) = 0\}$$

と定義すると, $S(X)$ は水準 $1-\alpha$ の信頼領域となる.

例 3.2 (母比率の近似信頼区間)
$X \sim B(n,\theta)$
　検定問題

$$H_0 : \theta = \theta_0 \text{ v.s. } H_1 : \theta \neq \theta_0$$

中心極限定理

$$Z = \frac{X - n\theta}{\sqrt{n\theta(1-\theta)}} \xrightarrow{d} N(0,1)$$

$$\frac{|X - n\theta_0|}{\sqrt{n\theta_0(1-\theta_0)}} > z_{\alpha/2} \quad \Rightarrow \quad H_0 \text{ を棄却}$$

$$\begin{aligned} S_x &= \left\{\theta_0;\ \frac{|x - n\theta_0|}{\sqrt{n\theta_0(1-\theta_0)}} \le z_{\alpha/2}\right\} \\ &= \left\{\theta_0;\ T_1(x) \le \theta_0 \le T_2(x)\right\}, \\ T_1(x) &= \frac{n}{n + z_{\alpha/2}^2}\left\{\hat{\theta} + \frac{z_{\alpha/2}^2}{2n} - \frac{z_{\alpha/2}}{\sqrt{n}}\sqrt{\hat{\theta}(1-\hat{\theta}) + \frac{z_{\alpha/2}^2}{4n}}\right\}, \\ T_2(x) &= \frac{n}{n + z_{\alpha/2}^2}\left\{\hat{\theta} + \frac{z_{\alpha/2}^2}{2n} + \frac{z_{\alpha/2}}{\sqrt{n}}\sqrt{\hat{\theta}(1-\hat{\theta}) + \frac{z_{\alpha/2}^2}{4n}}\right\}, \quad \hat{\theta} = \frac{x}{n} \end{aligned}$$

4　統計的決定問題

$X \sim P_\theta,\ \theta \in \Theta$
$\mathfrak{X}$: X の値域

点推定

$$\begin{aligned} &\delta : \mathfrak{X} \to \gamma(\Theta) \\ &\delta(X) : \gamma(\theta) \text{ の推定量} \end{aligned}$$

良さの基準

$$\begin{aligned} &\mathrm{E}_\theta[\{\delta(X) - \gamma(\theta)\}^2], \\ &\mathrm{E}_\theta[|\delta(X) - \gamma(\theta)|] \end{aligned}$$

仮説検定

$$\begin{aligned} &H_0 : \theta \in \Theta_0 \quad \text{v.s.} \quad H_1 : \theta \in \Theta_1 \quad (\Theta_0 \cap \Theta_1 = \emptyset), \\ &R : \text{ 検定の棄却域} \end{aligned}$$

$$\begin{aligned} &\delta : \mathfrak{X} \to \{H_0, H_1\}, \\ &\delta(x) = \begin{cases} H_0 & x \notin R \\ H_1 & x \in R \end{cases} \end{aligned}$$

良さの基準

$$\mathrm{P}_\theta(\delta(X) = H_1) \quad \theta \in \Theta_0, \quad \mathrm{P}_\theta(\delta(X) = H_0) \quad \theta \in \Theta_1$$

区間推定

$$\delta : \mathcal{X} \to \{[T_1, T_2] \subset \gamma(\Theta) \,|\, T_1 < T_2\}$$
$$\delta(X) = [T_1(X), T_2(X)] : \gamma(\theta) \text{ の信頼区間}$$

良さの基準

$$\mathrm{P}_\theta\{\theta \notin \delta(X)\},$$
$$\mathrm{E}_\theta[T_2(X) - T_1(X)]$$

定義 4.1 (統計的決定問題)
$X \sim P_\theta, \theta \in \Theta$
D : 可能な “戦略” の集合
δ : 決定関数, X の値域 $\mathcal{X}$ から D への写像
(X の実現値 x によって, 戦略 $\delta(x)$ を選択する.)
$L(\theta, d)$: 損失関数, $\Theta \times D$ 上の非負値関数
(未知母数 θ によって戦略の良しあしが数値で表現される)
期待損失 (リスク)

$$R(\theta, \delta) = \mathrm{E}_\theta[L(\theta, \delta(X))]$$

を最小とする決定関数を求めたい.

例 4.1
肝機能の診断
θ : 肝機能を表す特性値
x : 採血した血液中のある成分の量

$$\begin{cases} H_0 : & \theta < c_0 & \Rightarrow & \text{下方に異常} \\ H_1 : & c_0 \leq \theta \leq c_1 & \Rightarrow & \text{正常} \\ H_2 : & c_1 < \theta & \Rightarrow & \text{上方に異常} \end{cases}$$

＊. θ の値が同じでも, x の値が同じとは限らない.
θ の値が大きいとき, x の値も大きくなる傾向がある
問題 : x の値から, いかに H_0, H_1, H_2 を選択をするか
損失 : 各仮説 H_i と, 選択した仮説 H_j の食い違いで, 損失の大きさ $L_{i,j}(\theta)$ が決められる.

4.1 決定空間の要素

分布族 : $\mathcal{P} = \{P_\theta;\ \theta \in \Theta\}$
決定空間 : D
損失関数 : $L(\theta, d)$

分布族　2 項分布 $B(n, p)$, 正規分布 $N(\mu, \sigma^2)$, ポアソン分布 $P_o(\lambda)$, $\cdots$

例 4.2
$X = (X_1, \ldots, X_n)$: $X_1, \ldots, X_n \overset{i.i.d.}{\sim} P_\theta$

$$\theta = \begin{cases} \quad p & (2\text{ 項分布}) \\ (\mu, \sigma^2) & \text{正規分布}\lambda \quad \text{ポアソン分布} \end{cases}$$

$\gamma = \gamma(\theta) : \Theta \to \gamma(\Theta) \subset \mathbb{R}$

(i) $\gamma(\theta)$ が γ_0 より大きいかどうか判定したい

$$D = \{(d_0 : \gamma > \gamma_0), (d_1 : \gamma \leq \gamma_0)\}$$
$$L(\theta, d) = \begin{cases} 0 & \quad \text{正しい選択} \\ a & (d = d_0, \gamma(\theta) \leq \gamma_0) \\ b & (d = d_1, \gamma(\theta) > \gamma_0) \end{cases}$$
$$R(\theta, \delta) = \begin{cases} aP_\theta\{\delta(X) = d_0\} & (\gamma \leq \gamma_0) \\ bP_\theta\{\delta(X) = d_1\} & (\gamma > \gamma_0) \end{cases}$$

(ii) $\gamma(\theta)$ の点推定

$$D = \gamma(\Theta)$$
$$L(\gamma, d) = v(\gamma)w(|\gamma - d|), \quad v(\gamma) > 0, \quad w: \text{増加関数}$$
$$L(\gamma, d) = v(\gamma)(\gamma - d)^2, \quad v(\gamma) > 0$$

(iii) $\gamma(\theta)$ の区間推定

$$D = \{[l, u];\ -\infty < l < u < \infty\}$$
$$L_1(\theta, [l, u]) = \begin{cases} 0 & \gamma(\theta) \in [l, u] \\ w\left\{\left|\gamma(\theta) - \frac{1}{2}(u + l)\right|\right\} & \gamma(\theta) \notin [l, u] \end{cases}$$
$$w: \text{増加関数}$$
$$L_2(\theta, [l, u]) = l - u$$

2 つの損失関数に関する最適化問題

$$\mathrm{E}_\theta[L_1(\theta, [l, u])] \leq \alpha \text{ の条件下で } \mathrm{E}_\theta[L_2(\theta, [l, u])] \text{ を最小化する}$$

例 4.3 (判別問題)
$Z \sim B(1, \pi)$
$Z = z$ のときの X の条件付き分布の確率密度関数を $f_z(\mathrm{x})$
X の実現値から z の値を決めたい

$$D = \{z = 0, z = 1\}$$
$$L(z, d) = c_{zd}, \quad c_{00} = c_{11} = 0,\ c_{01} > 0, c_{10} > 0$$
$$R(\pi, \delta) = c_{01}\pi\mathrm{P}(\delta(X) = 1) + c_{10}(1 - \pi)\mathrm{P}(\delta(X) = 0)$$

例 4.4

$P_o(\lambda_1), P_o(\lambda_2)$: ポアソン分布, $\lambda_1 < \lambda_2$

$Z_1, \ldots, Z_n$: 独立な確率変数, $Z_i \sim P_o(\lambda_1)$ or $P_o(\lambda_2)$ $(i = 1, \ldots, n)$

$Z_i \sim P_o(\lambda_1)$ であるようなすべての i を選ぶ

$$D = \wp(\{1, 2, \ldots, n\})$$
$$L(\theta, d) = \#((d \cap d_0^c) \cup (d^c \cap d_0)), \quad \theta = (\lambda_1, \lambda_2, d_0)$$

ただし, d_0 は $Z_i \sim P_o(\lambda_1)$ であるようなすべての i からなる集合

最適化問題　リスク $R(\theta, \delta)$ を最小にする決定関数 δ を求める.

※. 一般に, θ ごとに リスクを最小とする決定関数は異なる.

各 θ に対して, $L(\theta, d) = 0$ となる $d = d_\theta$ が定まるとき,

θ_0 をひとつ決めて

$$\delta_0(x) \equiv d_{\theta_0}$$

と定めると, δ_0 は $\theta = \theta_0$ のとき $R(\theta, \delta)$ を最小とする.

2 つのアプローチ

1. 決定関数に制約条件を課して, δ_0 のような極端な決定関数を除外

    $\Rightarrow$ 不偏性と不変性

2. 決定関数に順序づけして, 必要最小限の決定関数の集合を特定する.

    $\Rightarrow$ 許容性

4.2　不偏性

定義 4.2 (不偏性)
δ : 決定関数
　L : 損失関数任意の $\theta, \theta' \in \Theta$ に対して

$$\mathrm{E}_\theta[L(\theta', \delta(X))] \geq \mathrm{E}_\theta[L(\theta, \delta(X))]$$

を満たすとき, δ は L-不偏であるという.

仮定 1 各 $\theta \in \Theta$ に対して, $L(\theta, d)$ を最小とする d がただ一つ存在する. この d を d_θ とする.

仮定 2 $d_{\theta_1} = d_{\theta_2} \Rightarrow$ 任意の d に対して $L(\theta_1, d) = L(\theta_2, d)$

このとき, $d' \in D$ に対して, $d' = d_\theta$ となる θ をとって

$$\ell(d', d) = L(\theta, d)$$

と定義すると, ℓ は, 正しい選択が d' であるときの d の損失を表わす.
　不偏性の条件は, 任意の θ, d' に対して

$$\mathrm{E}_\theta[\ell(d', \delta(X))] \geq \mathrm{E}_\theta[\ell(d_\theta, \delta(X))]$$

と表される. δ による選択は, 間違った選択よりも正しい選択に対してリスクが小さいことを意味する.

例 4.5 (例 4.2 の続き)

(i) $d_0 : \gamma(\theta) > \gamma_0,\ d_1 : \gamma(\theta) \leq \gamma_0$

$$\mathrm{E}_\theta[L(\theta', \delta(X))] = \begin{cases} a\mathrm{P}_\theta(\delta(X) = d_0) & (\gamma(\theta') \leq \gamma_0) \\ b\mathrm{P}_\theta(\delta(X) = d_1) & (\gamma(\theta') > \gamma_0) \end{cases}$$

不偏性の条件は

$$\begin{aligned} \gamma(\theta) > \gamma_0 \quad &\Rightarrow \quad a\mathrm{P}_\theta(\delta(X) = d_0) \geq b\mathrm{P}_\theta(\delta(X) = d_1) \\ \gamma(\theta) \leq \gamma_0 \quad &\Rightarrow \quad b\mathrm{P}_\theta(\delta(X) = d_1) \geq a\mathrm{P}_\theta(\delta(X) = d_0) \end{aligned}$$

$\mathrm{P}_\theta(\delta(X) = d_0) = 1 - \mathrm{P}_\theta(\delta(X) = d_1)$ より,

$$\begin{aligned} \gamma(\theta) > \gamma_0 \quad &\Rightarrow \quad \mathrm{P}_\theta(\delta(X) = d_1) \leq \frac{a}{a+b} \\ \gamma(\theta) \leq \gamma_0 \quad &\Rightarrow \quad \mathrm{P}_\theta(\delta(X) = d_1) \geq \frac{a}{a+b} \end{aligned}$$

d_0 を帰無仮説, d_1 を対立仮説とし, $\delta(X) = d_1$ を棄却することと考え, $\dfrac{a}{a+b}$ を有意水準と見ると, 不偏検定が得られる.

(ii) $\gamma(\theta) = \theta$ の点推定

$$L(\theta, d) = (\theta - d)^2$$

とすると, 不偏性の条件は, 任意の θ, θ' に対して

$$\mathrm{E}_\theta[(\theta' - \delta(X))^2] \geq \mathrm{E}_\theta[(\theta - \delta(X))^2].$$

$$\mathrm{E}_\theta[(\theta' - \delta(X))^2] = (\theta')^2 - 2\mathrm{E}_\theta[\delta(X)]\theta' + \{\mathrm{E}_\theta(\delta(X)^2)\}$$

が $\theta' = \mathrm{E}_\theta[\delta(X)]$ で最小となるので, 不偏性の条件は

$$\mathrm{E}_\theta[\delta(X)] = \theta$$

(iii) $\gamma(\theta) = \theta$ の信頼区間

$$L(\theta, [l, u]) = \begin{cases} 1 & \theta \notin [l, u] \\ 0 & \theta \in [l, u] \end{cases}$$

$\delta(X) = [L(X), U(X)]$ に対して

$$\mathrm{E}_\theta[L(\theta', \delta(X))] = \mathrm{P}_\theta(\theta' \notin [L(X), U(X)])$$

不偏性の条件は

$$\begin{aligned} &\mathrm{P}_\theta(\theta' \notin [L(X), U(X)]) \geq \mathrm{P}_\theta(\theta \notin [L(X), U(X)]) \\ &\Leftrightarrow \\ &\mathrm{P}_\theta(\theta' \in [L(X), U(X)]) \leq \mathrm{P}_\theta(\theta \in [L(X), U(X)]) \end{aligned}$$

4.3 許容性

定義 4.3 (決定関数の優劣)
δ, δ' : 決定関数

$$\begin{aligned} &{}^\forall\theta\, R(\theta, \delta) \leq R(\theta, \delta'), \\ &{}^\exists\theta\, R(\theta, \delta) < R(\theta, \delta'), \end{aligned}$$

が成り立つとき, δ は δ' に優るという.

定義 4.4 (許容性)
δ が許容的であるとは, δ に優る決定関数が存在しないこと

定義 4.5 (Bayes risk)
決定問題
　$\mathcal{P} = \{P_\theta \,|\, \theta \in \Theta\}$: 分布族
　D : 決定空間
　$L(\theta, d)$: 損失

に対して, θ を確率密度関数 $\rho(\theta)$ を持つ確率変数の実現値と考え,
P_θ を θ が与えられたときの条件付き分布とみなす.

$$\begin{aligned} r(\rho, \delta) &= \int \mathrm{E}_\theta[L(\theta, \delta(X))]\rho(\theta)d\theta \\ &= \int R(\theta, \delta)\rho(\theta)d\theta \end{aligned}$$

を ベイズリスクと呼ぶ. Bayes risk を最小とする決定関数をベイズ解と呼ぶ.

例 4.6 (判別問題, 例 4.3 の別表現)

$$\rho(\theta) : \mathrm{P}(\theta = 1) = \pi, \mathrm{P}(\theta = 0) = 1 - \pi$$
$$\theta = 0 \text{ のとき } X \sim f_0(x) \quad \theta = 1 \text{ のとき } X \sim f_1(x)$$
$$D = \{0, 1\},$$
$$L(\theta, d) = \begin{cases} 1 & (\theta \neq d) \\ 0 & (\theta = d) \end{cases} \quad (\theta, d = 0, 1)$$

決定関数 δ に対して

$$R_\delta = \{x;\ \delta(x) = 0\}$$

と定義して, δ と同一視する.

$$R(\theta, \delta) = \begin{cases} \displaystyle\int_{R_\delta} f_1(x)dx & \theta = 1 \\ \displaystyle\int_{R_\delta^c} f_0(x)dx & \theta = 0 \end{cases}$$
$$\begin{aligned} r(\rho, \delta) &= \pi \int_{R_\delta} f_1(x)dx + (1 - \pi) \int_{R_\delta^c} f_0(x)dx \\ &= (1 - \pi) + \int_{R_\delta} \{\pi f_1(x) - (1 - \pi)f_0(x)\}dx \end{aligned}$$

ベイズ解

$$R_\pi = \{x;\ \pi f_1(x) < (1 - \pi)f_0(x)\}$$

とすると, $r(\rho, R)$ は最小となる. この R_π をベイズ判別ルールと呼ぶ。
ベイズ判別ルールは許容的であることがわかるが, さらに任意の π に対して

$$\mathrm{P}(\pi f_1(X) = (1 - \pi)f_0(X) \,|\, X \sim f_j) = 0 \ (j = 0, 1)$$

ならば, 許容的判別ルールは, ある π に対するベイズ判別ルールに一致する.

例 4.7 (Stein の縮小推定量)
$\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n \overset{i.i.d.}{\sim} N_p(\boldsymbol{\mu}, \mathbf{I}_p), \quad \boldsymbol{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^n \boldsymbol{Y}_i$: 十分統計量

$$\boldsymbol{X} \sim N_p\Big(\boldsymbol{\mu}, \frac{1}{n}\mathbf{I}_p\Big),$$

$$
\begin{aligned}
&D = \{\boldsymbol{\mu};\ \boldsymbol{\mu} \in \mathbb{R}^p\},\\
&L(\boldsymbol{\mu}, \boldsymbol{d}) = (\boldsymbol{\mu} - \boldsymbol{d})^{\mathrm{T}}(\boldsymbol{\mu} - \boldsymbol{d})
\end{aligned}
$$

$\boldsymbol{X}$ は, $\boldsymbol{\mu}$ の一様最小分散不偏推定量であるが, $p \geq 3$ のとき許容的でない.

$$
\delta_a(\boldsymbol{X}) = \Big(1 - \frac{a}{\|\boldsymbol{X}\|^2}\Big)\boldsymbol{X}
$$

<u>リスク</u>

$$
R(\boldsymbol{\mu}, \delta_a) = \mathrm{E}[\|\boldsymbol{X} - \boldsymbol{\mu}\|^2] - 2a\mathrm{E}\Big[\frac{1}{\|\boldsymbol{X}\|^2}\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{X} - \boldsymbol{\mu})\Big] + a^2\mathrm{E}\Big[\frac{1}{\|\boldsymbol{X}\|^2}\Big]
$$

非心カイ 2 乗分布

$$
X_1, \ldots, X_m \ \ \text{：独立},\ X_i \sim N(\mu_i, 1)
$$

$Z = \displaystyle\sum_{i=1}^{m} X_i^2$ の分布を自由度 m, 非心度の $\delta = \displaystyle\sum_{i=1}^{m} \mu_i^2$ の非心カイ 2 乗分布といい, $\chi_m^2(\delta)$ と表す.

<u>特性関数</u>

$X = X_1, \mu = \mu_1$

$$
\begin{aligned}
\mathrm{E}[e^{itX^2}] &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{itx^2} e^{-\frac{1}{2}(x-\mu)^2} dx\\
&= (1 - 2it)^{-1/2} e^{-\mu^2/2} \exp\Big\{\frac{\mu^2}{2}\frac{1}{1 - 2it}\Big\}\\
\because\ \ (x - \mu)^2 - 2itx^2 &= (1 - 2it)x^2 - 2\mu x + \mu^2\\
&= (1 - 2it)\Big\{x - \frac{\mu}{1 - 2it}\Big\}^2 + \mu^2 - \frac{\mu^2}{1 - 2it}\\
&= (1 - 2it)\Big\{x - \frac{\mu}{1 - 2it}\Big\}^2 - \frac{2it}{1 - 2it}\mu^2\\
\mathrm{E}[e^{itZ}] &= (1 - 2it)^{-m/2} e^{-\delta/2} \exp\Big\{\frac{\delta}{2}\frac{1}{1 - 2it}\Big\}
\end{aligned}
$$

<u>確率密度関数</u>

$g_k(z)$ を χ_k^2 の確率密度関数とすると $\chi_m^2(\delta)$ の確率密度関数は

$$
g_m(z; \delta) = e^{-\delta/2} \sum_{k=0}^{\infty} \frac{(\delta/2)^k}{k!} g_{m+2k}(z)
$$

<u>モーメント</u>

$$
\begin{aligned}
&Z \sim \chi_k^2 : g_k(z) = \frac{1}{2\Gamma[k/2]} e^{-x/2}(x/2)^{k/2-1}\\
&\mathrm{E}[Z^r] = \frac{1}{2\Gamma[k/2]} \int_0^{\infty} e^{-x/2}(x/2)^{k/2-1} x^r dx\\
&\quad = \frac{2^r}{2\Gamma[k/2]} \int_0^{\infty} e^{-x/2}(x/2)^{r+k/2-1} dx = \frac{2^r\Gamma[r + k/2]}{\Gamma[k/2]} \quad (r + k > 0)
\end{aligned}
$$

$$
\begin{aligned}
&X \sim \chi_m(\delta) \\
&\mathrm{E}[X^{-1}] = e^{-\delta/2}\sum_{k=0}^{\infty}\frac{(\delta/2)^k}{k!}\frac{2^{-1}\Gamma[(m+2k-2)/2]}{\Gamma[(m+2k)/2]} \\
&\qquad = e^{-\delta/2}\sum_{k=0}^{\infty}\frac{(\delta/2)^k}{k!}\frac{1}{m+2k-2} \quad (m>2)
\end{aligned}
$$

$$
\boldsymbol{Z} = (Z_1,\ldots,Z_p)^{\mathrm{T}} = \sqrt{n}\mathbf{H}^{\mathrm{T}}\boldsymbol{X}, \quad \mathbf{H} = [\boldsymbol{h}_1,\ldots,\boldsymbol{h}_p], \quad \boldsymbol{h}_1 = \frac{1}{\|\boldsymbol{\mu}\|}\boldsymbol{\mu}
$$

$$
\Rightarrow \quad Z_1,\ldots,Z_p : \text{独立},\ Z_1 \sim N(\delta^{1/2},1),\ Z_2,\ldots,Z_p \sim N(0,1), \quad \delta^2 = n\|\boldsymbol{\mu}\|^2
$$

$$
\frac{1}{\|\boldsymbol{X}\|^2} = \frac{n}{\|\boldsymbol{Z}\|^2}
$$

$$
\frac{1}{\|\boldsymbol{X}\|^2}\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{X}-\boldsymbol{\mu}) = \frac{1}{\|\boldsymbol{Z}\|^2}\boldsymbol{Z}^{\mathrm{T}}(\boldsymbol{Z}-\sqrt{n}\mathbf{H}^{\mathrm{T}}\boldsymbol{\mu}) = 1 - \frac{\delta^{1/2}Z_1}{\sum_{i=1}^p Z_i^2}
$$

$$
\begin{aligned}
&\mathrm{E}\Big[\frac{1}{\sum_{i=1}^p Z_i^2}\Big] = (2\pi)^{-p/2}\int \frac{1}{\sum_{i=1}^p z_i^2}e^{-(z_1-\delta^{1/2})^2/2}e^{-(\sum_{i=2}^p z_i^2)/2}dz_1\cdots dz_p \\
&\delta^{1/2}\frac{\partial}{\partial\delta^{1/2}}\mathrm{E}\Big[\frac{1}{\sum_{i=1}^p Z_i^2}\Big] = \delta_1^{1/2}\mathrm{E}\Big[\frac{Z_1-\delta^{1/2}}{\sum_{i=1}^p Z_i^2}\Big] \\
&\mathrm{E}\Big[\frac{\delta^{1/2}Z_1}{\sum_{i=1}^p Z_i^2}\Big] = \mathrm{E}\Big[\frac{\delta}{\sum_{i=1}^p Z_i^2}\Big] + \delta^{1/2}\frac{\partial}{\partial\delta^{1/2}}\mathrm{E}\Big[\frac{1}{\sum_{i=1}^p Z_i^2}\Big]
\end{aligned}
$$

$$
\begin{aligned}
&\mathrm{E}\Big[\frac{1}{\sum_{i=1}^p Z_i^2}\Big] = e^{-\delta/2}\sum_{k=0}^{\infty}\frac{(\delta/2)^k}{k!}\frac{1}{p+2k-2} \\
&\delta^{1/2}\frac{\partial}{\partial\delta^{1/2}}\mathrm{E}\Big[\frac{1}{\sum_{i=1}^p Z_i^2}\Big] = -\delta e^{-\delta/2}\sum_{k=0}^{\infty}\frac{(\delta/2)^k}{k!}\frac{1}{p+2k-2} \\
&\qquad\qquad + \delta e^{-\delta/2}\sum_{k=0}^{\infty}\frac{(\delta/2)^{k-1}}{k!}\frac{k}{p+2k-2} \\
&1-\mathrm{E}\Big[\frac{\delta^{1/2}Z_1}{\sum_{i=1}^p Z_i^2}\Big] = 1 - e^{-\delta/2}\sum_{k=0}^{\infty}\frac{(\delta/2)^k}{k!}\frac{2k}{p+2k-2} \\
&\qquad = e^{-\delta/2}\sum_{k=0}^{\infty}\frac{(\delta/2)^k}{k!}\Big\{1-\frac{2k}{p+2k-2}\Big\} = \mathrm{E}\Big[\frac{p-2}{\sum_{i=1}^p Z_i^2}\Big]
\end{aligned}
$$

$$
\Rightarrow R(\boldsymbol{\mu},\delta_a) - R(\boldsymbol{\mu},\delta_0) = \mathrm{E}\Big[\frac{1}{\|\boldsymbol{X}\|^2}\Big]\frac{a^2-2a(p-2)}{n} \geq -\mathrm{E}\Big[\frac{1}{\|\boldsymbol{X}\|^2}\Big]\frac{(p-2)^2}{n}
$$

5　決定問題の不変性

5.1　決定問題の不変性

定義 5.1 (分布族の不変性)

$X \sim P_\theta \in \mathcal{P} = \{P_\theta;\ \theta \in \Theta\}$

$\mathfrak{X}$: X の値域
$G = \{g\}$: $\mathfrak{X}$ 上の変換群

$$^{\forall}\theta \in \Theta,\ ^{\forall}g \in G,\ ^{\exists!}\theta' \in \Theta,\ X \sim P_\theta \Rightarrow gX \sim P_{\theta'}$$

が成り立つとき, $\mathcal{P}$ は, G の下で不変であるという.
このとき, θ に θ' を対応させる Θ 上の変換を $\bar{g}$ と表す.

例 5.1 (位置尺度分布族)
$\mathcal{P} = \{P_\theta;\ \theta = (\sigma, \xi)' \in \Theta = (0, \infty) \times \mathbb{R}\}\}$

$$f(x;\theta) = \sigma^{-1} f((x-\xi)/\sigma) : P_\theta \text{ の pdf}$$
$$f(x) : \mathbb{R} \text{ 上で定義された pdf}$$

ξ を位置母数, σ を尺度母数という.
$G = \{g_{(a,b)};\ (a,b) \in (0,\infty) \times \mathbb{R}\}$

$$g_{(a,b)} : x \mapsto ax + b$$

G は合成変換を積として群をなす.

$$X \sim P_{(\sigma,\xi)} \Rightarrow g_{(a,b)}X = aX + b \equiv P_{(a\sigma, a\xi+b)}$$
$$\bar{g}_{(a,b)}(\sigma, \xi) = (a\sigma, a\xi + b)$$

※. $\bar{G} = \{\bar{g};\ g \in G\}$ は Θ 上の変換群である.

定義 5.2 (決定問題の不変性)
分布族 $\mathcal{P} = \{P_\theta;\ \theta \in \Theta\}$
決定空間 D
損失関数 $L(\theta, d)$
G : 観測変量 X の値域上の変換群
$\mathcal{P}$ が G の下で不変であり, Θ 上の変換 $\bar{g}$ に対応して D 上の変換 g^* が定まり, 写像

$$h : G \to G^*, \quad h(g) = g^*$$

$$h(g_1 g_2) = g_1^* g_2^*$$

を満たすとする. さらに, これら変換に関して損失関数が

$$L(\bar{g}\theta, g^* d) = L(\theta, d)$$

が成り立つとき, 決定問題は不変 (invariant) であるという.

例 5.2 (例 5.1 の続き位置母数の推定)
決定空間 $D = \mathbb{R}$

$$\bar{g}_{(a,b)}(\sigma, \xi) = (a\sigma, a\xi + b)$$

であるから

$$\bar{g}^*_{(a,b)}\xi = a\xi + b$$

損失関数

$$L_1((\sigma,\xi),d) = (\xi - d)^2 \ \text{不変ではない}$$
$$L_2((\sigma,\xi),d) = \Big(\frac{\xi - d}{\sigma}\Big)^2 \ \text{不変}$$

定義 5.3 (決定関数の不変性)
決定問題が不変であるとき,

$$^\forall x \in \mathfrak{X},\ ^\forall g \in G,\ \delta(gx) = g^*\delta(x)$$

を満たす決定関数を, G の下で共変であるという.
　また, $g \in G$ に対する決定空間上の変換 g^* が常に恒等変換になるとき, 共変な決定関数は不変であるという.

例 5.3
$x_1,\ldots,x_n$: ある物質の温度を n 回測定した結果 (摂氏)
　物質の真の温度 (摂氏) を μ, 測定誤差を独立に $N(0,\sigma^2)$ に従うと仮定すると.
　$x_1,\ldots,x_n$ は, $X_1,\ldots,X_n \overset{i.i.d.}{\sim} N(\mu,\sigma^2)$ の実現値と考えられる.
　$N(\mu,\sigma^2)$ は $f(x) = (2\pi)^{-1/2}\exp(-x^2/2)$ を基にした位置母数分布族で, μ の推定問題は, 例 5.2 の損失関数 L_2 を考えると, 例 5.1 の変換群 G の下で不変.

$$\delta(x_1,\ldots,x_n) = \frac{1}{n}\sum_{i=1}^n x_i$$

は共変推定量となる.
　測定値を華氏 で表したものを $y_1,\cdots,y_n$ とすると

$$y_i = \frac{9}{5}x_i + 32,$$
$$\delta(y_1,\cdots,y_n) = \frac{1}{n}\Big(\frac{9}{5}x_i + 32\Big) = \frac{9}{5}\delta(x_1,\ldots,x_n) + 32$$

なので, 華氏による推定値は, 摂氏で推定した結果を華氏に換算したものと一致する.

　※. $G^* = \{g^*;\ g \in G\}$ は D 上の変換群である.

補題 5.1 (リスクの不変性)
決定問題が G の下で不変であり, 決定関数 δ が G の下で共変ならば

$$R(\theta,\delta) = R(\bar{g}\theta,\delta)$$

が成り立つ. したがって, 共変な決定関数のリスクは $\bar{G}$ による Θ 上の軌道上定数.
　特に, $\bar{G}$ の Θ への作用が推移的ならば δ のリスクは未知母数の値に依存しない。

証明

$$
\begin{aligned}
R(\theta,\delta) &= \mathrm{E}_\theta[L(\theta,\delta(X))] = \mathrm{E}_\theta[L(\bar{g}\theta, g^*\delta(X))] \\
&= \mathrm{E}_\theta[L(\bar{g}\theta,\delta(gX))] = \mathrm{E}_{\bar{g}\theta}[L(\bar{g}\theta,\delta(X))] \\
&= R(\bar{g}\theta,\delta)
\end{aligned}
$$

■

5.2 分散の共変推定量

決定問題

$$
\begin{aligned}
&\boldsymbol{X} = (X_1,\cdots,X_n)', \quad X_1,\cdots,X_n \overset{i.i.d.}{\sim} N(\mu,\sigma^2) \\
&D = \{\sigma^2 | \sigma \in (0,\infty)\} \quad (\sigma^2 \text{ の推定問題 }) \\
&L((\mu,\sigma),d) = \Big(\frac{d-\sigma^2}{\sigma^2}\Big)^2
\end{aligned}
$$

変換群

$$
G = \{g_{a,b} | g_{a,b}(\boldsymbol{x}) = (\tfrac{x_1-a}{b},\cdots,\tfrac{x_n-a}{b}), a\in\mathbb{R}, b>0\}
$$

とすると、決定問題は不変で、母数空間上の変換、決定空間上の変換は、それぞれ、

$$
\bar{g}_{a,b}(\mu,\sigma) = \Big(\frac{\mu-a}{b},\frac{\sigma}{b}\Big), \quad g^*_{a,b}(d) = \frac{d}{b^2}
$$

共変推定量

δ を σ^2 の共変推定量とすると

$$
\delta\Big(\frac{X_1-a}{b},\cdots,\frac{X_n-a}{b}\Big) = \frac{1}{b^2}\delta(X_1,\cdots,X_n)
$$

$$
a = \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i, \quad b = B_n = \sqrt{\sum_{i=1}^n (X_i-\bar{X})^2}
$$

ととると

$$
\delta(X_1,\cdots,X_n) = B_n^2\delta\Big(\frac{X_1-\bar{X}}{B_n},\cdots,\frac{X_n-\bar{X}}{B_n}\Big)
$$

逆に、任意の n 変数関数 $\gamma(y_1,\cdots,y_n)$ を用いて

$$
\delta(X_1,\cdots,X_n) = B_n^2\gamma\Big(\frac{X_1-\bar{X}}{B_n},\cdots,\frac{X_n-\bar{X}}{B_n}\Big)
$$

と定義すると、δ は共変推定量となる。

共変推定量の分布について

$$
\begin{aligned}
&(X_1-\bar{X},\cdots,X_n-\bar{X})'=\Pi_n\boldsymbol{X},\\
&\quad \Pi_n=I_n-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n',\\
&\quad \mathbf{1}_n=(1,1,\cdots,1)':n\times 1\ \text{ベクトル}\\
&\quad I_n:n\times n\ \text{単位行列}\ ,
\end{aligned}
$$

と表されるが、Π_n の固有値は、1 ($(n-1)$ 重根) と 0 (単根) なので、固有値 1 に属する一次独立な $n-1$ 個の長さ 1 の固有ベクトルを横に並べてできる $(n\times(n-1))$-行列を H とすると、

$$\Pi_n=HH',\quad H'H=I_{n-1}$$

が成り立つ。($(\mu,\sigma)=(0,1)$) のとき、$\boldsymbol{X}\sim N_n(\mathbf{0},I_n)$ となるので、

$$\boldsymbol{Y}=H'\boldsymbol{X}\sim N_{n-1}(\mathbf{0},I_{n-1})$$

となり、$B_n=\|\boldsymbol{Y}\|$ となるので、

$$
\begin{aligned}
&B_n^2\sim\chi^2_{n-1}\\
&\quad pdf:g(x)=\frac{1}{2^{n-1}\Gamma[\frac{n-1}{2}]}e^{-x/2}\big(\frac{x}{2}\big)^{\frac{n-1}{2}-1}\quad(x>0)\\
&\quad Y_1=B_n\cos\Theta_1,\\
&\quad Y_2=B_n\sin\Theta_1\cos\Theta_2,\\
&\quad Y_3=B_n\sin\Theta_1\sin\Theta_2\cos\Theta_3,\\
&\qquad\qquad\vdots\\
&\quad Y_{n-2}=B_n(\sin\Theta_1\sin\Theta_2\cdots\sin\Theta_{n-3})\cos\Theta_{n-2},\\
&\quad Y_{n-1}=B_n(\sin\Theta_1\sin\Theta_2\cdots\sin\Theta_{n-3})\sin\Theta_{n-2},\\
&\quad 0\le\Theta_i\le\pi\ (i=1,\cdots,n-3);0\le\Theta_{n-2}<2\pi
\end{aligned}
$$

と曲座標変換すると、B_n と $(\Theta_1,\cdots,\Theta_{n-2})$ が独立であることがわかる。

(ヤコビアンは $B_n^{n-2}\displaystyle\prod_{j=1}^{n-3}\sin^{n-2-j}\Theta_j$)

$\dfrac{X_i-\bar{X}}{B_n}$ $(i=1,\cdots,n)$ は、$\Theta_1,\cdots,\Theta_{n-2}$ の関数なので、

$$B_n^2\ \text{と}\ C:=\gamma\Big(\frac{X_1-\bar{X}}{B_n},\cdots,\frac{X_n-\bar{X}}{B_n}\Big)\ \text{は独立}$$

最小リスク共変推定量

リスク（期待損失）の不変性から

$$
\begin{aligned}
&\mathrm{E}_{(\mu,\sigma)}[L((\mu,\sigma),B_n^2C)]=\mathrm{E}_{(0,1)}[L((0,1),B_n^2C)]\\
&\quad=\mathrm{E}[(B_n^2C-1)^2]=\mathrm{E}[\mathrm{E}[B_n^4]C^2-2\mathrm{E}[B_n^2]C+1]\\
&\quad=\mathrm{E}[(n-1)(n+1)C^2-2(n-1)C+1]
\end{aligned}
$$

$$C\equiv\frac{1}{n+1}\quad\Big(\delta(X_1,\ldots,X_n)=\frac{1}{n+1}\sum_{i=1}^{n}(X_i-\bar{X})^2\Big)$$

ととると、期待損失は最小となる。

5.3 不変検定

定義 5.4 (検定問題の不変性)

分布族 $\mathcal{P} = \{P_\theta;\ \theta \in \Theta\}$: 変換群 G の下で不変

決定空間 $D = \{H_0 : \theta \in \Theta_0, H_1 : \theta \in \Theta_1\}$

$(\Theta_0, \Theta_1 \subset \Theta, \Theta_0 \cap \Theta_1 = \emptyset)$

損失関数

$$L(\theta, d) = \begin{cases} 1 & (\theta \in \Theta_0, d = H_1) \\ 1 & (\theta \in \Theta_1, d = H_0) \\ 0 & \text{(その他)} \end{cases}$$

$${}^\forall g \in G,\ \bar{g}(\Theta_0) = \Theta_0, \bar{g}(\Theta_1) = \Theta_1,$$

が成り立つとき, 検定問題は不変であるという.

$\bar{g}\theta \in \Theta_0$ と $\theta \in \Theta_0$ は同値であるので

$$g^* d = d$$

決定空間上の変換 g^* はすべて恒等変換となるので, 共変な決定関数（検定関数）は不変検定と呼ばれ, 検定関数 $\phi(x)$ が不変であるための条件は

$$\phi(gx) = \phi(x)$$

確率化検定の決定空間　検定関数 : $0 \le \phi(x) \le 1$, 確率 $\phi(x)$ で H_0 を棄却する.

このとき

決定空間 : $D = [0, 1]$ (閉区間)

損失は, $U \sim U(0, 1)$ として

$$L(\theta, d) = \begin{cases} \mathrm{P}_\theta(U \le d) & (\theta \in \Theta_0) \\ \mathrm{P}_\theta(U > d) & (\theta \in \Theta_1) \end{cases}$$

検定関数 ϕ のリスクは

$$R(\theta, \phi) = \begin{cases} \mathrm{E}_\theta[\mathrm{P}_\theta(U \le \phi(X)|X)] = \mathrm{P}_\theta(U \ge \phi(X)) & (\theta \in \Theta_0) \\ \mathrm{E}_\theta[\mathrm{P}_\theta(U > \phi(X)|X)] = \mathrm{P}_\theta(U > \phi(X)) & (\theta \in \Theta_1) \end{cases}$$

確率化検定に対しても

$$\phi(gx) = \phi(x)$$

を満たすとき, 不変検定と呼ぶ.

5.4 母平均の同等性に関する不変検定

検定問題

$$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu_x, \sigma^2), \quad Y_1, \ldots, Y_n \overset{i.i.d.}{\sim} N(\mu_y, \sigma^2)$$

$$\text{帰無仮説 } H_0 : \mu_x = \mu_y \text{ v.s. 対立仮説 } H_1 : \mu_x \ne \mu_y$$

十分統計量

$$\theta = (\mu_x, \mu_y, \sigma^2)$$

$$f(x_1, \ldots, x_n, y_1, \ldots, y_n;\ \theta) = \frac{1}{(2\pi\sigma^2)^n} \exp\Big[-\frac{1}{2\sigma^2} \sum_{i=1}^n \{(x_i - \mu_x)^2 + (y_i - \mu_y)^2\}\Big]$$

$$\sum_{i=1}^n \{(x_i - \mu_x)^2 + (y_i - \mu_y)^2\} = \sum_{i=1}^n \{(x_i - \bar{x})^2 + (y_i - \bar{y})^2\} + n\{(\bar{x} - \mu_x)^2 + (\bar{y} - \mu_y)^2\},$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^n y_i$$

より $T = (\bar{X}, \bar{Y}, S^2)$ は十分統計量:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i, \quad \bar{Y} = \frac{1}{n}\sum_{i=1}^n Y_i, \quad S^2 = \sum_{i=1}^n \{(X_i - \bar{X})^2 + (Y_i - \bar{Y})^2\}$$

$$\bar{X} \sim N(\mu_x, \frac{\sigma^2}{n}), \quad \bar{Y} \sim N(\mu_y, \frac{\sigma^2}{n}), \quad \frac{S^2}{\sigma^2} \sim \chi^2_{2(n-1)}$$

$\phi(x_1, \ldots, x_n, y_1, \ldots, y_n)$ を有意水準 α の検定関数

$$\begin{aligned} &\phi^T(t) = \mathrm{E}[\phi(X_1, \ldots, X_n, Y_1, \ldots, Y_n) \,|\, T = t] \\ \Rightarrow \quad &\mathrm{E}[\phi^T(T)] = \mathrm{E}[\phi(X_1, \ldots, X_n, Y_1, \ldots, Y_n)] \end{aligned}$$

より, $\phi^T(t)$ も有意水準 α の検定関数で, 検出力も ϕ と同じ.
※ $\phi^T(t)$ は, $0, 1$ 以外の値をとる可能性があり, 一般に確率化検定となる.

変換群

$$G = \{g_{a,b} \,|\, a \neq 0, b \in \mathbb{R}\}$$
$$g_{a,b}(X_1, \ldots, X_n, Y_1, \ldots, Y_n) = (aX_1 + b, \ldots, aX_n + b, aY_1 + b, \ldots, aY_n + b)$$
$$aX_i + b \sim N(a\mu_x + b, a^2\sigma^2), \quad aY_i + b \sim N(a\mu_y + b, a^2\sigma^2),$$

$\Rightarrow \bar{g}_{a,b}(\mu_x, \mu_y, \sigma^2) = (a\mu_x + b, a\mu_y + b, a^2\sigma^2)$ パラメータの変換

$$g_{a,b}(X_1, \ldots, X_n, Y_1, \ldots, Y_n) = (\tilde{X}_1, \ldots, \tilde{X}_n, \tilde{Y}_1, \ldots, \tilde{Y}_n)$$
$$\bar{\tilde{X}} := \frac{1}{n}\sum_{i=1}^n \tilde{X}_i = a\bar{X} + b, \quad \bar{\tilde{Y}} := \frac{1}{n}\sum_{i=1}^n \tilde{Y}_i = a\bar{Y} + b,$$
$$\tilde{S^2} := \sum_{i=1}^n \{(\tilde{X}_i - \bar{\tilde{X}})^2 + (\tilde{Y}_i - \bar{\tilde{Y}})^2\} = a^2 S^2$$

$g^T_{a,b}(\bar{X}, \bar{Y}, S^2) = (a\bar{X} + b, a\bar{Y} + b, a^2 S^2)$: 十分統計量の変換

十分統計量を用いた不変検定　$\phi^T(\bar{x}, \bar{y}, s^2)$ を十分統計量による不変検定とすると

$$\forall a, \forall b,\ \phi^T(\bar{x}, \bar{y}, s^2) = \phi^T(a\bar{x} + b, a\bar{y} + b, a^2 s^2)$$

$a = \dfrac{\mathrm{sign}(\bar{x}-\bar{y})}{s}, b = -a\bar{y}$ ととると

$$\phi^T(\bar{x}, \bar{y}, s^2) = \phi^T\Big(\frac{|\bar{x}-\bar{y}|}{s}, 0, 1\Big)$$

したがって, $\phi^T(\bar{x}, \bar{y}, s^2)$ が不変検定であるための必要十分条件は,
$\phi^T(\bar{x}, \bar{y}, s^2)$ が $f = \dfrac{n(\bar{x}-\bar{y})^2}{2s^2}$ の関数であることである.

$F = \dfrac{n(\bar{X}-\bar{Y})^2}{S^2}$ **の分布**

$$Z = \frac{\sqrt{n}(\bar{X}-\bar{Y})}{\sqrt{2}\sigma} \sim N(\delta, 1), \quad \delta = \frac{\sqrt{n}(\mu_x - \mu_y)}{\sqrt{2}\sigma}$$

であるから, $W = Z^2 \sim \chi^2_1(\delta^2)$ であり, 確率密度関数は

$$f_W(w; \delta^2) = \sum_{k=0}^{\infty} e^{-\delta^2/2} \frac{(\delta^2/2)^k}{k!} \frac{1}{2^{k+1/2}\Gamma[k+\frac{1}{2}]} e^{-(w/2)} w^{k-1/2}$$

$V = \dfrac{S^2}{\sigma^2} \sim \chi^2_{2n-2}$ だから V の確率密度関数は

$$f_V(v) = \frac{1}{2^{n-1}\Gamma[n-1]} e^{-v/2} v^{n-2}$$

$$F = \frac{W}{V}, \quad \begin{cases} w = fv \\ v = v \end{cases}, \quad \begin{vmatrix} \dfrac{\partial w}{\partial f} & \dfrac{\partial w}{\partial v} \\ \dfrac{\partial v}{\partial f} & \dfrac{\partial v}{\partial v} \end{vmatrix} = \begin{vmatrix} v & f \\ 0 & 1 \end{vmatrix} = v$$

$$\begin{aligned} f_{F,V}(f, v; \delta^2) &= \sum_{k=0}^{\infty} e^{-\delta^2/2} \frac{(\delta^2/2)^k}{k!} \frac{1}{2^{k+1/2}\Gamma[k+\frac{1}{2}]} e^{-(fv/2)} (fv)^{k-1/2} \frac{1}{2^{n-1}\Gamma[n-1]} e^{-v/2} v^{n-2} v \\ &= \sum_{k=0}^{\infty} e^{-\delta^2/2} \frac{(\delta^2/2)^k}{k!} \frac{1}{2^{n+k-1/2}\Gamma[n-1]\Gamma[k+\frac{1}{2}]} f^{k-1/2} e^{-(1+f)v/2)} v^{n+k-3/2} \end{aligned}$$

$$\begin{aligned} f_F(f; \delta^2) &= \int_0^{\infty} f_{F,V}(f, v) dv \\ &= \sum_{k=0}^{\infty} e^{-\delta^2/2} \frac{(\delta^2/2)^k}{k!} \frac{\Gamma[n+k-\frac{1}{2}]}{\Gamma[n-1]\Gamma[k+\frac{1}{2}]} f^{k-1/2} \frac{1}{(1+f)^{n+k-1/2}} \\ &= f_F(f; 0) g(f; \delta^2), \end{aligned}$$

$$f_F(f; 0) = \frac{\Gamma[n-\frac{1}{2}]}{\Gamma[n-1]\Gamma[\frac{1}{2}]} f^{-1/2} \frac{1}{(1+f)^{n-1/2}}$$

$$g(f; \delta^2) = \sum_{k=0}^{\infty} e^{-\delta^2/2} \frac{(\delta^2/2)^k}{k!} \frac{\Gamma[n+k-\frac{1}{2}]\sqrt{\pi}}{\Gamma[n-\frac{1}{2}]\Gamma[k+\frac{1}{2}]} \Big(\frac{f}{1+f}\Big)^k$$

一様最強力不変検定　不変検定は F に基づく検定

$$\mu_x = \mu_y \Leftrightarrow \delta^2 = 0$$

$$H_0 : \delta^2 = 0 \quad \text{v.s} \quad H_1 : \delta^2 = d_1 > 0$$

ネイマンピアソンの基本定理より, H_0, H_1 に対する最強力検定は

$$\begin{aligned} &\frac{f_F(f;\delta_1^2)}{f_F(f;0)} = g(f;\delta_1^2) > c \quad \Rightarrow \quad H_0 \text{ を棄却} \\ &\Leftrightarrow \quad f > c' \quad \Rightarrow \quad H_0 \text{ を棄却} \end{aligned} \tag{5.8}$$

有意水準 α に対して, c' は

$$\int_{c'}^{\infty} f_F(f;0)df = \alpha$$

とすればよい. (5.8) は δ_1^2 の値に依存しないから対立仮説 $H_a : \delta^2 > 0$ に対して一様最強力検定である.

5.5 位置母数の Pitman 推定量

(Lehmann の点推定の第 3 章より抜粋)

定義 5.5
(位置母数の共変推定量) 分布族 $\mathcal{P} = \{P_\theta;\ \theta \in \mathbb{R}\}$
$f(x_1, \dots, x_n, \theta)$: P_θ の pdf

変換群 $G = \{g_a;\ a \in \mathbb{R}\}$

$$g_a(x_1, \dots, x_n) = (x_1 + a, \dots, x_n + a)$$

G の下で $\mathcal{P}$ が不変であるためには、

$$f(x_1, \dots, x_n;\ \theta) = g(x_1 - \theta, \dots, x_n - \theta)$$

$$\bar{g}_a(\theta) = \theta + a$$

θ の推定量 $\delta(x_1, \dots, x_n)$ が共変であるためには

$$\delta(x_1 + a, \dots, x_n + a) = \delta(x_1, \dots, x_n) + a$$

この δ を (位置) 共変推定量という.(location equivariant estimator)
損失関数 $L(\theta, d)$ が不変であるためには

$$L(\theta, d) = \rho(d - \theta)$$

補題 5.2
共変推定量のリスク、偏差、分散は θ に依存しない.

定義 5.6 (最小リスク共変推定量 (MRE:Minimum Risk Equivariant estimator))
共変推定量の中で, リスクを最小とする推定量をいう.

定理 5.1
δ_0 を任意の共変推定量とすると, δ が共変推定量であるための必要十分条件は, 可測関数 $v(y_1, \ldots, v_{n-1})$ が存在して

$$\delta(x_1, \ldots, x_n) = \delta_0(x_1, \cdot, x_n) - v(x_1 - x_n, \ldots, x_{n-1} - x_n)$$

と書けることである.

定理 5.2 (MRE)
$y_i = x_i - x_n$ $(i = 1, \ldots, n-1)$ とする. リスクが有限な共変推定量 δ_0 が存在し, 各 $y = (y_1, \ldots, y_{n-1})$ に対して

$$\mathrm{E}_{\theta=0}[\rho(\delta_0(X) - v)|y]$$

を最小とする $v = v^*(y)$ が存在するならば,

$$\delta^*(x) = \delta_0(x) - v(y)$$

は MRE である.

系 5.3
定理 5.2 の仮定の下で, ρ が単調でない凸関数ならば, MRE は存在し, 狭義の凸関数ならば, MRE は一意に存在する.

系 5.4
定理 5.2 の仮定の下で, $\rho(d - \theta) = (d - \theta)^2$ ならば

$$v^*(y) = \mathrm{E}_{\theta=0}[\delta_0(X)|y]$$

例 5.4 (正規分布)
$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, σ^2 : 既知
ρ が凸の偶関数ならば, $\bar{X}$ は MRE である.

定理 5.5 (正規分布の "least faboroble" 性)
$\mathcal{F} = \{F\}$ を, 分散が 1 である 1 変量分布の全体からなる分布族とし, f をその pdf とする.

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} f(x_i - \theta)$ とし, $r_n(F)$ を その 2 乗損失に関する MRE のリスクとすると, $r_n(F)$ は $F = N(0, 1)$ のとき最大となる.

定理 5.6 (位置母数の Pitman 推定量)
定理 5.2 の仮定の下で, 2 乗損失に関する MRE は

$$\delta^*(x) = \frac{\int_{-\infty}^{\infty} u\ g(x_1 - u, \ldots, x_n - u)du}{\int_{-\infty}^{\infty} g(x_1 - u, \ldots, x_n - u)du}$$

によって与えられる.

例 5.5
一様分布 $X_1, \ldots, X_n \overset{i.i.d.}{\sim} U(\theta - b/2, \theta + b/2)$, b : 既知

$$g(x_1 - \theta, \ldots, x_n - \theta) = \begin{cases} b^{-n} & \theta - \dfrac{b}{2} \leq x_{(1)} \leq x_{(n)} \leq \theta + \dfrac{b}{n} \\ 0 & \text{その他} \end{cases}$$

$$\delta^*(x) = \frac{1}{2}(X_{(1)} + X_{(n)})$$