数理解析研究所講究録
916 巻 1995 年 131-148

Selecting the number of components in a mixture of normal distributions: A simple case

大阪府立大学 大学院 工学研究科　初山光司　(Mitsuji Hatsuyama)
筑波大学 数学系　狩野　裕　(Yutaka Kano)
大阪府立大学 工学部　長尾壽夫　(Hisao Nagao)

1 序文

本稿では、観測データが単一分布モデルに従っているのか混合分布モデルに従っているのか、という問題を考える。例えば、混合分布モデルのコンポーネント数が２個の場合は以下のようにかける。

$$\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n \overset{i.i.d.}{\sim} P(\boldsymbol{x}|\pi, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \pi f(\boldsymbol{x}|\boldsymbol{\theta}_1) + (1-\pi) f(\boldsymbol{x}|\boldsymbol{\theta}_2)$$

ここで、$f(\boldsymbol{x}|\boldsymbol{\theta})$ は確率密度関数であり、π $(0 \leq \pi \leq 1)$ は混合比率である。

このとき、帰無仮説は単一分布モデルに従っている、対立仮説は混合分布モデルに従っているとすると、各仮説は以下のようになる。

$$\begin{aligned} H_0 &: \pi = 1 \\ H_1 &: 0 \leq \pi < 1 \end{aligned}$$

この問題に対して尤度比検定法 (ＬＲＴ) を適用することが考えられるが、通常の漸近 χ^2 近似が成り立たないことが知られており、また、尤度比検定統計量は無限大に発散し、最尤推定量が存在しない場合もある (Hartigan(1985))。本稿２節で、ＬＲＴの問題点を述べ、ＬＲＴが適用できる場合とできない場合を紹介する。本稿３節で、ＬＲＴ以外に提案されている４つの方法 (Bayes 法など) を紹介する。しかし、この方法では、パラメータの一致推定量が得られるのみであり漸近分布を理論的に求めることは困難である。

そこで、本稿４節では、２つの正規分布による混合分布モデルに対し、モーメント法に基づく検定法を提案する。この方法を用いることによって、従来の方法では導出が困難であった推定量の漸近分布を得ることができ、この漸近分布を利用して、第１種の過誤と第２種の過誤を理論的に評価できる。また、この方法の数値的検討も行った。

2 尤度比検定とその問題点

問題を簡略化するため、2つの1変量正規分布の混合分布 ($m=2$) を取り上げ尤度比検定の問題点を検討する。以下のような状況を考える。

$$X_1, X_2, \cdots, X_n \overset{i.i.d.}{\sim} P(x|\pi, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = \pi N(x|\mu_1, \sigma_1^2) + (1-\pi) N(x|\mu_2, \sigma_2^2)$$

とし、仮説を以下のように設定する。

$$\begin{aligned} H_0 &: \pi = 1 \\ H_1 &: 0 \leq \pi < 1 \end{aligned}$$

この仮説に対する尤度比検定統計量 (LRT) は、

$$\lambda_n = \frac{\max\limits_{H_0} \prod\limits_{i=1}^{n} P(X_i|\pi, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2)}{\max\limits_{H_0 \cup H_1} \prod\limits_{i=1}^{n} P(X_i|\pi, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2)}$$

で表される。これに通常の漸近 χ^2 近似を形式的に適用すると（自由度は各仮説の下での自由なパラメータ数の差である）

$$-2\log\lambda_n \xrightarrow{\mathcal{L}} \chi_3^2$$

となる。しかしながら、種々の正則条件が満足されず、この χ^2 近似は正しくない。

通常の漸近 χ^2 近似が正しくない理由は、以下の3点であると思われる。

1. $identifiability$の欠如の問題。
2. H_0の下での真値$\pi=1$はパラメータ空間の内点ではない。
3. H_0の下で情報行列が特異となっていることがある。

本節では、これらの問題点と、問題2のような場合における尤度比検定の方法 (Chernoff(1953))、ならびにその漸近分布である $\bar{\chi}^2$-分布 (Shapiro(1985,88)) を紹介する。

2.1 尤度比検定が適用できない場合

$$X_1, X_2, \cdots, X_n \overset{i.i.d.}{\sim} P(x|\pi, \mu) = \pi N(x|0,1) + (1-\pi) N(x|\mu, 1)$$

とし、仮説を以下のようにおく。

$$\begin{aligned} H_0 &: \pi = 1 \\ H_1 &: 0 \leq \pi < 1 \end{aligned}$$

このとき、Hartigan(1985) より、

$$-2\log\lambda_n \xrightarrow{P} \infty \quad as \quad n \longrightarrow \infty$$

が成立する。

2.2 尤度比検定が適用できる場合

2.2.1 Chernoff(1953)-Shapiro(1985,88) の手法

p 次元観測データが、互いに独立に確率分布 $f(\boldsymbol{x}|\boldsymbol{\theta})$ をもつ同一分布に従っているとする。

$$\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n \overset{i.i.d.}{\sim} f(\boldsymbol{x}|\boldsymbol{\theta})$$

仮説を以下のようにおく。

$$\begin{aligned} H_0 &: \boldsymbol{\theta} \in \boldsymbol{\omega} \\ H_1 &: \boldsymbol{\theta} \in \boldsymbol{\tau} \end{aligned}$$

このとき、真値 $\boldsymbol{\theta}_0$ が $\boldsymbol{\omega}$ の境界にある場合の尤度比検定統計量の漸近分布を求める (Chernoff(1953))。そのために、いくつかの定義を述べる。

（1）

$$\text{集合}\boldsymbol{C} \subset \boldsymbol{R}^q : \text{凸錐} \Longleftrightarrow \quad \begin{aligned} &\boldsymbol{C} : \text{凸} \\ &a \geq 0, \quad \boldsymbol{\theta} \in \boldsymbol{C} \Rightarrow a\boldsymbol{\theta} \in \boldsymbol{C} \end{aligned}$$

（2）

$\boldsymbol{C}_{\boldsymbol{\phi}}$: $\boldsymbol{x}_0$ における ϕ の近似凸錐 $\Longleftrightarrow$

$$\begin{aligned} \inf_{\boldsymbol{y} \in \boldsymbol{C}_{\boldsymbol{\phi}}} \|(\boldsymbol{x} - \boldsymbol{x}_0) - \boldsymbol{y}\| &= o(\|\boldsymbol{x} - \boldsymbol{x}_0\|) \quad for \quad \boldsymbol{x} \in \phi \\ \inf_{\boldsymbol{x} \in \boldsymbol{\phi}} \|(\boldsymbol{x} - \boldsymbol{x}_0) - \boldsymbol{y}\| &= o(\|\boldsymbol{y}\|) \quad for \quad \boldsymbol{y} \in \boldsymbol{C}_{\boldsymbol{\phi}} \end{aligned}$$

（3）

$\boldsymbol{X} \sim N_q(0, J^{-1})$ であり、$\boldsymbol{C}$ は q次元凸錐とするとき、$\bar{\chi}^2$ を以下のように定義する。

$$\bar{\chi}^2 = \inf_{\boldsymbol{\theta} \in \boldsymbol{C}} (\boldsymbol{X} - \boldsymbol{\theta})' J (\boldsymbol{X} - \boldsymbol{\theta})$$

$\bar{\chi}^2$ の分布は χ^2 分布の重みつきの和つまり、

$$P\{\bar{\chi}^2 \geq c^2\} = \sum_{i=0}^{q} w_i P\{\chi_i^2 \geq c^2\}$$

と表される。ここで、w_i は $\sum_{i=0}^{q} w_i = 1,\ w_i \geq 0, i = 1, \cdots, q$ という重みで $\chi_0^2 = 0$ とする。

Chernoff(1953) の結果

$$\begin{aligned} -2\log\lambda &= -2\log \frac{\max_{\boldsymbol{\theta} \in \boldsymbol{\omega}} \prod_{i=1}^{n} f(\boldsymbol{X}_i|\boldsymbol{\theta})}{\max_{\boldsymbol{\theta} \in \boldsymbol{\tau}} \prod_{i=1}^{n} f(\boldsymbol{X}_i|\boldsymbol{\theta})} \\ &\xrightarrow{\mathcal{L}} \min_{\boldsymbol{\theta} \in \boldsymbol{C}_{\boldsymbol{\omega}}} (\boldsymbol{Z} - \boldsymbol{\theta})' J (\boldsymbol{Z} - \boldsymbol{\theta}) - \min_{\boldsymbol{\theta} \in \boldsymbol{C}_{\boldsymbol{\tau}}} (\boldsymbol{Z} - \boldsymbol{\theta})' J (\boldsymbol{Z} - \boldsymbol{\theta}) \end{aligned}$$

ただし、$Z \sim N_q(0, J^{-1})$ である。
Shapiro(1985,88) の結果

$$\boldsymbol{C_\omega} \subset \boldsymbol{C_\tau}$$
$$\boldsymbol{C_\omega} \text{と} \boldsymbol{C_\tau} \text{のどちらかが線形空間}$$

であるとき

$$\min_{\boldsymbol{\theta} \in \boldsymbol{C_\omega}} (\boldsymbol{Z} - \boldsymbol{\theta})' J (\boldsymbol{Z} - \boldsymbol{\theta}) - \min_{\boldsymbol{\theta} \in \boldsymbol{C_\tau}} (\boldsymbol{Z} - \boldsymbol{\theta})' J (\boldsymbol{Z} - \boldsymbol{\theta}) = \min_{\boldsymbol{\theta} \in (\boldsymbol{C}^*)^o} (\boldsymbol{Z} - \boldsymbol{\theta})' J (\boldsymbol{Z} - \boldsymbol{\theta})$$

ただし、

$$\begin{aligned} \boldsymbol{C}^* &= \boldsymbol{C_\tau} \cap (\boldsymbol{C_\omega})^\perp \\ \boldsymbol{C}^o &= \{\boldsymbol{y} : (\boldsymbol{x}, \boldsymbol{y}) \le 0 \ for \ all \ \boldsymbol{x} \in \boldsymbol{C}\} \ (\boldsymbol{C}^o \text{は} \boldsymbol{C} \text{の双対空間}) \end{aligned}$$

次節において、この Chernoff(1953)-Shapiro(1985,88) による $\bar{\chi}^2$ 分布の手法を、混合分布モデルにおけるコンポーネント数の選定に適用する。

2.2.2　コンポーネント数が2つで分布が既知の場合

ここで、以下のような仮説に対して尤度比検定を行う。

$$\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n \overset{i.i.d.}{\sim} P(\boldsymbol{x}|\pi) = \pi f(\boldsymbol{x}) + (1-\pi) g(\boldsymbol{x}) \quad , \ 0 \le \pi \le 1$$

$$\begin{aligned} H_0 &: \ \pi = 1 \\ H_1 &: \ 0 \le \pi < 1 \end{aligned}$$

ただし、密度関数 $f, g (f \neq g)$ は既知とする。

$$\lambda = \frac{\max_{\pi=1} \prod_{i=1}^{n} P(\boldsymbol{X}_i|\pi)}{\max_{0 \le \pi < 1} \prod_{i=1}^{n} P(\boldsymbol{X}_i|\pi)}$$

とおく。このとき、$\boldsymbol{\omega} = \{1\}, \boldsymbol{\tau} = [0,1]$ であるから、Chernoff(1953) より近似凸錐は $\boldsymbol{C_\omega} = \{0\}, \boldsymbol{C_\tau} = (-\infty, 0]$ とかける。よって、

$$\begin{aligned} -2 \log \lambda \ &\xrightarrow{\mathcal{L}} \ \inf_{\theta \in \boldsymbol{C_\omega}} (Z - \theta)^2 J - \inf_{\theta \in \boldsymbol{C_\tau}} (Z - \theta)^2 J, \qquad Z \sim N(0, J^{-1}) \\ &= \ \inf_{\theta = 0} (Z - \theta)^2 - \inf_{\theta \le 0} (Z - \theta)^2, \qquad Z \sim N(0, 1) \end{aligned}$$

となる。ここで、

$$\begin{aligned} \boldsymbol{C_\omega} &= \{\theta : \theta = 0\} \\ \boldsymbol{C_\tau} &= \{\theta : \theta \le 0\} \end{aligned}$$

これより

$$C^* = \boldsymbol{C_{\mathcal{T}}} \cap \boldsymbol{C_{\omega}^{\perp}} = \{\theta : \theta \le 0\}$$
$$(C^*)^o = \{\theta : \theta \ge 0\}$$

このとき Shapiro(1985) により、

$$\begin{aligned} -2\log\lambda &\xrightarrow{\mathcal{L}} \min_{\boldsymbol{\theta}\in \boldsymbol{C_{\omega}}} (Z-\theta)^2 - \min_{\boldsymbol{\theta}\in \boldsymbol{C_{\mathcal{T}}}} (Z-\theta)^2 \\ &= \min_{\boldsymbol{\theta}\in (C^*)^o} (Z-\theta)^2 = \min_{\theta\ge 0}(Z-\theta)^2 \end{aligned}$$

である。$(Z \sim N(0,1))$

また、漸近分布は

$$\begin{aligned} P\{-2\log\lambda \le c^2\} &\longrightarrow P\{\min_{\theta\ge 0}(Z-\theta)^2 \le c^2\} \\ &= \frac{1}{2}P\{0 \le c^2\} + \frac{1}{2}P\{Z^2 \le c^2 | Z \le 0\} \end{aligned}$$

ただし、$Z \sim N(0,1)$ である。これより

$$-2\log\lambda \xrightarrow{\mathcal{L}} \frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$$

となる（$\chi_0^2 = 0$）。ここで、

$$\begin{aligned} 1-\alpha &= \frac{1}{2}P(0 < c^2) + \frac{1}{2}P(\chi_1^2 < c^2) \\ 2\alpha &= P(\chi_1^2 \ge c^2) \end{aligned}$$

であるから、有意水準 α の棄却域は $-2\log\lambda \ge \chi_1^2(2\alpha)$ となる。

2.2.3 コンポーネント数が 3 つ以上で分布が既知の場合

ここで、以下のような仮説に対して尤度比検定を行う。

$$\begin{aligned} \boldsymbol{X}_1, \cdots, \boldsymbol{X}_n &\overset{i.i.d.}{\sim} \sum_{i=1}^{m} \pi_i f_i(\boldsymbol{x}) \\ 0 \le \pi_i \le 1, &\qquad \sum_{i=1}^{m} \pi_i = 1 \end{aligned}$$

コンポーネント数が 3 つの時

$$-2\log\lambda \xrightarrow{\mathcal{L}} \frac{1}{6}\chi_0^2 + \frac{1}{2}\chi_1^2 + \frac{1}{3}\chi_2^2$$

コンポーネント数が 4 つの時

$$-2\log\lambda \xrightarrow{\mathcal{L}} \frac{1}{36}\chi_0^2 + \frac{1}{4}\chi_1^2 + \frac{17}{36}\chi_2^2 + \frac{1}{4}\chi_3^2$$

H_0 の下での真値 $\pi = 1$ はパラメータ空間の境界点であるため、漸近 χ^2 近似は正しくなく、$\bar{\chi}^2$ 分布となる。

3　その他の方法: Bayes 法など

ここでは、Henna(1985),Leroux(1992),Chen and Kalbfleisch(1993) らにより与えられた有限混合分布モデルのパラメータ（コンポーネント数 m を含む）の一致推定量及び、Aitkin and Rubin(1985) により与えられたコンポーネント数の事前分布を用いたEMアルゴリズムによるパラメータの推定とその検定を紹介する。しかし、これらの方法では、統計量の漸近分布が求まらないので、統計的仮説検定は行えない。

定義と記号

$\boldsymbol{x} \in \boldsymbol{R}^p,\ \boldsymbol{\theta} \in \Theta \subset \boldsymbol{R}^q$

$\{f(\boldsymbol{x},\boldsymbol{\theta})|\boldsymbol{\theta} \in \Theta\}$: 密度の族

$\{F(\boldsymbol{x},\boldsymbol{\theta})|\boldsymbol{\theta} \in \Theta\}$: 分布の族

$G(\boldsymbol{\theta})$: Θ上の分布関数

$f(\boldsymbol{x},G) = \int_\Theta f(\boldsymbol{x},\boldsymbol{\theta})G(d\boldsymbol{\theta})$: 混合分布モデルの密度関数

$G_m(\boldsymbol{\theta}) = \sum_{i=1}^{m} \pi_i I(\boldsymbol{\theta}_i \le \boldsymbol{\theta})$: Θ上の離散分布関数（$\sum_{i=1}^{m} \pi_i = 1, 0 \le \pi \le 1, \boldsymbol{\theta}_i \in \Theta$）

$f(\boldsymbol{x},G_m) = \int_\Theta f(\boldsymbol{x},\boldsymbol{\theta})G_m(d\boldsymbol{\theta}) = \sum_{i=1}^{m} \pi_i f(\boldsymbol{x},\boldsymbol{\theta}_i)$: 有限混合モデルの密度関数

$F(\boldsymbol{x},G_m) = \sum_{i=1}^{m} \pi_i F(\boldsymbol{x},\boldsymbol{\theta}_i)$: 有限混合モデルの分布関数

$\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n \overset{i.i.d.}{\sim} f(\boldsymbol{x},G_m)$

$F_n(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} I(\boldsymbol{X}_i \le \boldsymbol{x})$: 経験分布関数

$d(F_1,F_2)$: 2つの分布関数 F_1,F_2間の距離関数

m^0 : 真のコンポーネント数

$G^0_{m^0}(\boldsymbol{\theta})$: 真の離散分布関数

（1）　Henna(1985): Minimum Distance

$$T_n^H(G_m) = \int \{F(\boldsymbol{x},G_m) - F_n(\boldsymbol{x})\}^2 F_n(d\boldsymbol{x})$$

とおく。また、$\lambda(n)$ は以下の条件を満たす実数列とする。

$$\lambda(n) \longrightarrow \infty,\ \frac{\lambda^2(n)}{n} \longrightarrow 0\ (n \to \infty),$$

$$\sum_{n=1}^{\infty} \frac{\lambda^2(n)}{n} \exp\{-2\lambda^2(n)\} < \infty$$

このとき、m の推定量 $\hat{m}$ を以下の最小化問題の解として定める。

$$\hat{m} = \min_{m}\{m \in \boldsymbol{N} | \min_{\pi_i, \boldsymbol{\theta}_i} T_n^H(G_m) < \frac{\lambda^2(n)}{n}\}$$

例えば (Chung(1949)) は $\frac{\lambda^2(n)}{n} = \lambda(n, \epsilon) = \frac{(1+\epsilon)}{2}(\log\log n)/n \ ({}^\forall\epsilon > 0)$ としている。

（2） Leroux(1992): Penalized Maximum Likelihood

$$T_n^L(G_m) = -\int \log f(\boldsymbol{x}, G_m)F_n(d\boldsymbol{x}) + a_{m,n}/n$$

とおく。ここで、$a_{m,n}$ は以下の条件を満たす実数列である。

$$a_{m,n} > 0, a_{m+1,n} > a_{m,n}, a_{m,n}/n \longrightarrow 0\,(n \to \infty)$$

このとき、m の推定量$\hat{m}$を

$$\min_{m}\{\min_{\pi_i, \boldsymbol{\theta}_i} T_n^L(G_m)\}$$

の最小化問題の解として定める。また、$T_n^L(G_{\hat{m}})$ の最小値を与える $G_{\hat{m}}$ を $\hat{G}_{\hat{m}}$ とかく。

（3） Chen and Kalbfleisch(1993): Penalized Minimum Distance

$$T_n^c(G_m) = d(F_n(\boldsymbol{x}), F(\boldsymbol{x}, G_m)) - c_n \sum_{i=1}^{m} \log \pi_i$$

とおく。ここで、c_n は以下の条件を満たす実数列である。

$$d(F_n(\boldsymbol{x}), F(\boldsymbol{x}, G_{m^0}^0)) = o_p(c_n), c_n = o(1), c_n > 0$$

このとき、m の推定量 $\hat{m}$ を

$$\min_{m}\{\min_{\pi_i, \boldsymbol{\theta}_i} T_n^c(G_m)\}$$

の最小化問題の解として定める。また、$T_n^c(G_{\hat{m}})$ の最小値を与える $G_{\hat{m}}$ を $\hat{G}_{\hat{m}}$ とかく。

（4） Aitkin and Rubin(1985): Bayes 法

$k = 1, \cdots, M-1 \quad M : given$ として以下の検定を逐次繰り返すことにより m の推定を行う。

$$\begin{aligned} H_0 &: m = k \\ H_1 &: m = M \end{aligned}$$

このとき、検定統計量は以下のようである。

$$-2\log \frac{\max_{\boldsymbol{\theta}_j, j=1,\cdots,k} \int \prod_{i=1}^n f(\boldsymbol{X}_i, G_k)p_k(\boldsymbol{\pi})d\boldsymbol{\pi}}{\max_{\boldsymbol{\theta}_j, j=1,\cdots,M} \int \prod_{i=1}^n f(\boldsymbol{X}_i, G_M)p_M(\boldsymbol{\pi})d\boldsymbol{\pi}}$$

ここで、$p_m(\boldsymbol{\pi})$ は $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_m)'$ の事前分布である。Aitkin and Rubin(1985) は、事前分布の下で $\boldsymbol{X}_i$ 間の相関はあまり大きくないので、この統計量の漸近分布は自由度 $(M-k)dim(\boldsymbol{\theta})$ の χ^2 分布としてよいと述べている。

4　モーメント法に基づく検定法

この節では、観測データが1変量の場合に以下の仮説に対し、モーメント法でパラメータを推定し、その漸近分布を求める。

$$X_1, X_2, \cdots, X_n \overset{i.i.d.}{\sim} P(x|\pi, \mu_1, \mu_2, \sigma^2) = \pi N(x|\mu_1, \sigma^2) + (1-\pi)N(x|\mu_2, \sigma^2)$$

$$H_0 : \pi = 1$$
$$H_1 : 0 \leq \pi < 1$$

モーメント法は計算が易しく考え方も平易であり、従来の方法では導出が困難であった推定量の漸近分布を得ることができる。この漸近分布を利用して、第1種の過誤（正しくは $m = 1$ であるのに $m = 2$ としてしまう誤り）と第2種の過誤（正しくは $m = 2$ であるのに $m = 1$ としてしまう誤り）を理論的に評価することができる。以下この方法に必要な定義や計算を行う。

$P(x|\pi, \mu_1, \mu_2, \sigma^2)$ の積率母関数は、以下のようである。

$$g(\theta) \;=\; \pi \exp(\mu_1\theta + \frac{1}{2}\sigma^2\theta^2) + (1-\pi)\exp(\mu_2\theta + \frac{1}{2}\sigma^2\theta^2)$$

よって、

$$\begin{aligned}
E(X\ \ ) = g'(0) &= \pi\mu_1 + (1-\pi)\mu_2 \\
E(X^2) = g''(0) &= \pi(\mu_1^2 + \sigma^2) + (1-\pi)(\mu_2^2 + \sigma^2) \\
E(X^3) = g'''(0) &= \pi(3\sigma^2\mu_1 + \mu_1^3) + (1-\pi)(3\sigma^2\mu_2 + \mu_2^3) \\
E(X^4) = g''''(0) &= \pi(3\sigma^4 + 6\sigma^2\mu_1^2 + \mu_1^4) + (1-\pi)(3\sigma^4 + 6\sigma^2\mu_2^2 + \mu_2^4)
\end{aligned}$$

ここで、以下のようにおき、$\hat{\pi}, \hat{\mu_1}, \hat{\mu_2}, \hat{\sigma^2}$ を次の方程式の解とする。

$$\tilde{X} \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i = \pi\mu_1 + (1-\pi)\mu_2 \tag{4.0.1}$$

$$\tilde{Y} \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \pi(\mu_1^2 + \sigma^2) + (1-\pi)(\mu_2^2 + \sigma^2) \tag{4.0.2}$$

$$\tilde{Z} \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i^3 = \pi(3\sigma^2\mu_1 + \mu_1^3) + (1-\pi)(3\sigma^2\mu_2 + \mu_2^3) \tag{4.0.3}$$

$$\tilde{W} \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i^4 = \pi(3\sigma^4 + 6\sigma^2\mu_1^2 + \mu_1^4) + (1-\pi)(3\sigma^4 + 6\sigma^2\mu_2^2 + \mu_2^4) \tag{4.0.4}$$

また、次に以下のように定義すると、

$$X \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i = \tilde{X}$$

$$
\begin{aligned}
Y &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \tilde{X})^2 = \tilde{Y} - \tilde{X}^2 \\
Z &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \tilde{X})^3 = \tilde{Z} - 3\tilde{X}\tilde{Y} + 2\tilde{X}^3 \\
W &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \tilde{X})^4 = \tilde{W} - 4\tilde{X}\tilde{Z} + 6\tilde{X}^2\tilde{Y} - 3\tilde{X}^4
\end{aligned}
$$

上記の推定方程式 (4.0.1),(4.0.2),(4.0.3),(4.0.4) は、以下のようにかける。

$$
\begin{aligned}
X &= (1-\pi)(\mu_2 - \mu_1) + \mu_1 && (4.0.5) \\
Y &= \pi(1-\pi)(\mu_2 - \mu_1)^2 + \sigma^2 && (4.0.6) \\
Z &= \pi(1-\pi)(1-2\pi)(\mu_1 - \mu_2)^3 && (4.0.7) \\
W &= \pi(1-\pi)(1-3\pi+3\pi^2)(\mu_2 - \mu_1)^4 \\
&\quad +6\pi(1-\pi)\sigma^2(\mu_2 - \mu_1)^2 + 3\sigma^4 && (4.0.8) \\
W - 3Y^2 &= \pi(1-\pi)(1-6\pi+6\pi^2)(\mu_2 - \mu_1)^4 && (4.0.9)
\end{aligned}
$$

中心極限定理より、これらの統計量の漸近分布は同時 4 次元多変量正規分布に従う。

$$
\begin{aligned}
\sqrt{n}(X - \mu_1) &\xrightarrow{\mathcal{L}} N(0, \sigma^2) \\
\sqrt{n}(Y - \sigma^2) &\xrightarrow{\mathcal{L}} N(0, 2\sigma^4) \\
\sqrt{n}Z &\xrightarrow{\mathcal{L}} N(0, 15\sigma^6) \\
\sqrt{n}(W - 3Y^2) &\xrightarrow{\mathcal{L}} N(0, 24\sigma^8)
\end{aligned}
$$

また、$\sqrt{n}(X-\mu_1)$ と $\sqrt{n}(Y-\sigma^2)$、$\sqrt{n}(Y-\sigma^2)$ と $\sqrt{n}Z$、$\sqrt{n}(W-3\sigma^2)$ と $\sqrt{n}Z$ は漸近独立となる。

ここで種々の状況の下で$\hat{\pi}$を構成しよう。

4.1 μ_1, σ^2が既知, 未知パラメータがπ, μ_2の 2 つの場合

推定方程式は、以下のようにかける。

$$
\begin{aligned}
X &= (1-\pi)(\mu_2 - \mu_1) + \mu_1 && (4.1.10) \\
Y &= \pi(1-\pi)(\mu_2 - \mu_1)^2 + \sigma^2 && (4.1.11)
\end{aligned}
$$

この推定方程式より μ_2 を消去すると、

$$
\pi(X - \mu_1)^2 = (1-\pi)(Y - \sigma^2)
$$

となり、π について解くと、

$$
\hat{\pi} = \frac{Y - \sigma^2}{(X - \mu_1)^2 + (Y - \sigma^2)}
$$

が得られる。これより H_0 の下、

$$\begin{aligned}\hat{\pi}-1 &= \frac{-(X-\mu_1)^2}{(X-\mu_1)^2+(Y-\sigma^2)} \\ \sqrt{n}(\hat{\pi}-1) &= \frac{-\{\sqrt{n}(X-\mu_1)\}^2}{\sqrt{n}(Y-\sigma^2)}+O_p(\frac{1}{\sqrt{n}}) \end{aligned} \tag{4.1.12}$$

となる。ここで $H_0:\pi=1$ の下で $\sqrt{n}(X-\mu_1), \sqrt{n}(Y-\sigma^2)$ は独立でそれぞれ $N(0,\sigma^2), N(0,2\sigma^4)$ に分布収束するので、(4.1.12) より、$\hat{\pi}$ の H_0 の下での漸近分布は以下のようになる。

$$-\sqrt{n}(\hat{\pi}-1) \xrightarrow{\mathcal{L}} \frac{N^2(0,\sigma^2)}{N(0,2\sigma^4)}$$

つまり、

$$-\sqrt{2n}(\hat{\pi}-1) \xrightarrow{\mathcal{L}} \frac{N^2(0,1)}{N(0,1)} \tag{4.1.13}$$

となる。

次に、$0\le\pi\le 1$ という条件を考慮する場合について考える。つまり

$$\hat{\pi}^* = \begin{cases} 1 & 1\le\hat{\pi} \\ \hat{\pi} & 0<\hat{\pi}<1 \\ 0 & \hat{\pi}\le 0 \end{cases}$$

とおき、推定量 $\hat{\pi}^*$ の漸近分布を求める。

$$\begin{aligned} P\{-\sqrt{2n}(\hat{\pi}^*-1)<x\} &= P\{-\sqrt{2n}(\hat{\pi}^*-1)<x,\ 1\le\hat{\pi}\} \\ &+ P\{-\sqrt{2n}(\hat{\pi}^*-1)<x,\ 0<\hat{\pi}<1\} \\ &+ P\{-\sqrt{2n}(\hat{\pi}^*-1)<x,\ \hat{\pi}\le 0\} \end{aligned} \tag{4.1.14}$$

ここで、$x>0$ である。また

$$\begin{aligned} P\{\hat{\pi}\le 0\} &= P\{-\sqrt{2}\frac{\{\sqrt{n}(X-\mu_1)\}^2}{\sqrt{n}(Y-\sigma^2)}\le -\sqrt{2n}\}+o(1)\to 0 \quad (n\to\infty) \\ P\{\hat{\pi}\ge 1\} &= P\{-\sqrt{2}\frac{\{\sqrt{n}(X-\mu_1)\}^2}{\sqrt{n}(Y-\sigma^2)}\ge 0\}+o(1) \\ &= P\{\sqrt{n}(Y-\sigma^2)\le 0\}+o(1)\to\frac{1}{2} \quad (n\to\infty) \end{aligned}$$

である。これより (4.1.14) は、

$$\begin{aligned} \text{第1項} \quad &= P\{-\sqrt{2n}(\hat{\pi}^*-1)<x,\ \hat{\pi}^*=1\} = P\{0<x,\ \hat{\pi}^*=1\} \\ &= P\{\hat{\pi}^*=1\} = P\{\hat{\pi}\ge 1\} \xrightarrow{\mathcal{P}} \frac{1}{2} \end{aligned} \tag{4.1.15}$$

$$\text{第3項} \quad \le P\{\hat{\pi}\le 0\} \xrightarrow{\mathcal{P}} 0 \tag{4.1.16}$$

また、推定方程式 (4.1.11) より $0<\hat{\pi}<1 \Leftrightarrow 0<Y-\sigma^2$ であるから

$$\begin{aligned}
\text{第 2 項} \quad &= P\{-\sqrt{2n}(\hat{\pi}-1)<x,\ Y-\sigma^2>0\} \\
&= P\{\sqrt{2}\frac{\{\sqrt{n}(X-\mu_1)\}^2}{\sqrt{n}(Y-\sigma^2)}<x,\ Y-\sigma^2>0\}+o(1) \\
&= P\{\sqrt{2}\frac{\{\sqrt{n}(X-\mu_1)\}^2}{x}<\sqrt{n}(Y-\sigma^2),\ Y-\sigma^2>0\}+o(1) \\
&= P\{\sqrt{2}\frac{\{\sqrt{n}(X-\mu_1)\}^2}{x}<\sqrt{n}(Y-\sigma^2)\}+o(1) \\
&= \frac{1}{2}P\{\sqrt{2}\frac{\{\sqrt{n}(X-\mu_1)\}^2}{x}<\sqrt{n}|Y-\sigma^2|\}+o(1) \\
&= \frac{1}{2}P\{\sqrt{2}\frac{\{\sqrt{n}(X-\mu_1)\}^2}{\sqrt{n}|Y-\sigma^2|}<x\}+o(1) \qquad (4.1.17)
\end{aligned}$$

となり、

$$P\{-\sqrt{2n}(\hat{\pi}^*-1)<x\} \xrightarrow{\mathcal{P}} \frac{1}{2}+\frac{1}{2}P\{\frac{N^2(0,1)}{|N(0,1)|}<x\} \qquad (4.1.18)$$

が得られる。

この統計量 $-\sqrt{2n}(\hat{\pi}^*-1)$ を用いて混合分布モデルの仮説検定が行える。ここで、$x>0$ とし、

$$\begin{aligned}
1-\alpha &= \frac{1}{2}+\frac{1}{2}P(\frac{N^2(0,1)}{|N(0,1)|}<x) \\
2\alpha &= P(\frac{N^2(0,1)}{|N(0,1)|}\geq x) \\
\alpha &= P(\frac{N^2(0,1)}{N(0,1)}\geq x)
\end{aligned}$$

であるから、有意水準 α の検定関数 $\varphi(\boldsymbol{X})$ は

$$\varphi(\boldsymbol{X}) = \begin{cases} 1 & -\sqrt{2n}(\hat{\pi}^*-1) > \dfrac{N^2(0,1)}{N(0,1)}(\alpha) \\ 0 & -\sqrt{2n}(\hat{\pi}^*-1) \leq \dfrac{N^2(0,1)}{N(0,1)}(\alpha) \end{cases}$$

と与えられる。ただし、$\frac{N^2(0,1)}{N(0,1)}(\alpha)$ は $\frac{N^2(0,1)}{N(0,1)}$ 分布の上側 α 点とする。

この検定は、$\hat{\pi}$ と $\hat{\pi}^*$ で漸近的に同等であることがわかる。

次に、この検定の理論的検出力を求めるため、この統計量 $-\sqrt{2n}(\hat{\pi}-1)$ の H_1 の下での漸近分布を考える。デルタ法により、

$$-\sqrt{2n}(\hat{\pi}-\pi) \xrightarrow{\mathcal{L}} N(0,2\sigma_{\hat{\pi}}^2) \qquad (4.1.19)$$

が得られる。$-\sqrt{n}(\hat{\pi}-\pi)$ の漸近分散は、$d=(\mu_2-\mu_1)/\sigma$ とおくと、

$$\sigma_{\hat{\pi}}^2 = \pi(1-\pi)+\frac{4\pi}{d^2}+\frac{2}{d^4}+o(1)$$

と表される。以上より $-\sqrt{2n}(\hat{\pi}-1)$ の H_1 の下での分布を、

$$N(-\sqrt{2n}(\pi-1), 2\sigma_{\hat{\pi}}^2) \tag{4.1.20}$$

で近似しよう。このとき

(1) データ数 n の増加に対して、検出力は増加する。

(2) π の増減に対する検出力の増減は、$\sigma_{\hat{\pi}}^2$ が π に依存しているため一概にいえないが、n が十分の大きければ、π の減少に伴い（H_0 から遠くなる）検出力は増加する。

(3) 漸近分散 $\sigma_{\hat{\pi}}^2$ は d の減少関数である。これは d の増加に伴い $\hat{\pi}$ の推定の精度が向上するということを意味する。

(4) (3) は、ただちに検出力の増加を意味しないが、n が大きく

$$-\sqrt{2n}(\pi-1) > \frac{N^2(0,1)}{N(0,1)}(\alpha)$$

であるとき、d の増加に伴い（H_0 から遠くなる）検出力は増加する。

この理論的検出力の数値例、およびコンピュータシュミレーションの結果は後で述べる。

4.2 σ^2が既知, 未知パラメータがπ, μ_1, μ_2の 3 つの場合

推定方程式は、以下のようにかける。

$$X = (1-\pi)(\mu_2-\mu_1)+\mu_1 \tag{4.2.21}$$
$$Y = \pi(1-\pi)(\mu_2-\mu_1)^2+\sigma^2 \tag{4.2.22}$$
$$Z = \pi(1-\pi)(1-2\pi)(\mu_1-\mu_2)^3 \tag{4.2.23}$$

この推定方程式より $\hat{\pi}$ は、以下の式を満たすことがわかる。

$$(1-2\pi)^2(Y-\sigma^2)^3 = \pi(1-\pi)Z^2 \tag{4.2.24}$$

これは、π の推定値が 2 つ存在する可能性があることを意味している。実際、(4.2.24) は $\pi=\frac{1}{2}$ で対称であり、$\hat{\pi}=\pi$ が解であるなら $\hat{\pi}=1-\pi$ も解である。これは、混合分布モデルにおける identifiability の欠如のためである。（パラメータが $(\pi,\mu_1,\mu_2,\sigma^2)$ のときと $(1-\pi,\mu_2,\mu_1,\sigma^2)$ のときとでは $P(x|\pi,\mu_1,\mu_2,\sigma^2)$ は同じ分布となってしまう。）この問題を避けるため、π を $\frac{1}{2}\le\pi\le 1$ に制限して考えることにする。つまり、(4.2.24) の 2 解のうち 1 に近い値を与える解を $\hat{\pi}$ とするということである。このとき、$\hat{\pi}$ の H_0 の下での漸近分布を求める。

(4.2.24) より、

$$\begin{aligned}\hat{\pi} &= \frac{4(Y-\sigma^2)^3+Z^2+\sqrt{4(Y-\sigma^2)^3Z^2+Z^4}}{2\{4(Y-\sigma^2)^3+Z^2\}} \\ \hat{\pi}-1 &= \frac{-4(Y-\sigma^2)^3-Z^2+\sqrt{4(Y-\sigma^2)^3Z^2+Z^4}}{2\{4(Y-\sigma^2)^3+Z^2\}}\end{aligned} \tag{4.2.25}$$

ととる。ここでマクローリン展開を $\sqrt{4(Y-\sigma^2)^3Z^2+Z^4}$ に適用すると

$$\sqrt{4(Y-\sigma^2)^3Z^2+Z^4}=Z^2+2(Y-\sigma^2)^3+o_p(n^{-\frac{3}{2}})$$

となり

$$\begin{aligned}\hat{\pi}-1 &= \frac{-2(Y-\sigma^2)^3+o_p(n^{-\frac{3}{2}})}{2\{4(Y-\sigma^2)^3+Z^2\}} \\ \sqrt{n}(\hat{\pi}-1) &= \frac{-2\{\sqrt{n}(Y-\sigma^2)\}^3+o_p(1)}{2(\sqrt{n}Z)^2+O_p(n^{-\frac{1}{2}})} \\ &= \frac{-\{\sqrt{n}(Y-\sigma^2)\}^3}{(\sqrt{n}Z)^2}+O_p(n^{-\frac{1}{2}})\end{aligned} \tag{4.2.26}$$

となり、H_0 の下での漸近分布は以下のようになる。

$$-\sqrt{n}(\hat{\pi}-1) \xrightarrow{\mathcal{L}} \frac{N^3(0,2\sigma^4)}{N^2(0,15\sigma^6)}$$

つまり、

$$-\frac{15}{2\sqrt{2}}\sqrt{n}(\hat{\pi}-1) \xrightarrow{\mathcal{L}} \frac{N^3(0,1)}{N^2(0,1)} \tag{4.2.27}$$

となる。

また、$0 \leq \pi \leq 1$ という制約をつけた場合の $\hat{\pi}$ の漸近分布は、先ほどと同様にすると、

$$P\{-\frac{15}{2\sqrt{2}}\sqrt{n}(\hat{\pi}^*-1)<x\} \xrightarrow{\mathcal{P}} \frac{1}{2}+\frac{1}{2}P\{\frac{|N^3(0,1)|}{N^2(0,1)}<x\} \tag{4.2.28}$$

が得られる。

次に、この検定の理論的検出力を求めるため、この統計量 $-\frac{15}{2\sqrt{2}}\sqrt{n}(\hat{\pi}-1)$ の H_1 の下での漸近分布を考える。デルタ法により、

$$-\frac{15}{2\sqrt{2}}\sqrt{n}(\hat{\pi}-\pi) \xrightarrow{\mathcal{L}} N(0,\frac{225}{8}\sigma_{\hat{\pi}}^2) \tag{4.2.29}$$

が得られる。ここで、$d=(\mu_2-\mu_1)/\sigma$ とおくと、$-\sqrt{n}(\hat{\pi}-\pi)$ の漸近分散は、

$$\begin{aligned}\sigma_{\hat{\pi}}^2 &= \pi(1-\pi)(1-6\pi+6\pi^2)^2-\frac{12\pi(1-\pi)(1-9\pi+9\pi^2)}{d^2} \\ &+\frac{18(1+6\pi-6\pi^2)}{d^4}+\frac{60}{d^6}+o(1)\end{aligned}$$

と表される。以上より $-\frac{15}{2\sqrt{2}}\sqrt{n}(\hat{\pi}-1)$ の H_1 の下での分布は、

$$N(-\frac{15}{2\sqrt{2}}\sqrt{n}(\pi-1), \frac{225}{8}\sigma_{\hat{\pi}}^2) \tag{4.2.30}$$

で近似される。

4.3 未知パラメータが$\pi, \mu_1, \mu_2, \sigma^2$の4つの場合

推定方程式は、以下のようにかける。

$$\begin{aligned}
X &= (1-\pi)(\mu_2-\mu_1)+\mu_1 && (4.3.31)\\
Y &= \pi(1-\pi)(\mu_2-\mu_1)^2+\sigma^2 && (4.3.32)\\
Z &= \pi(1-\pi)(1-2\pi)(\mu_1-\mu_2)^3 && (4.3.33)\\
W &= \pi(1-\pi)(1-3\pi+3\pi^2)(\mu_2-\mu_1)^4+6\pi(1-\pi)\sigma^2(\mu_2-\mu_1)^2+3\sigma^4 \\
&= \pi(1-\pi)(1-6\pi+6\pi^2)(\mu_2-\mu_1)^4+3Y^2 && (4.3.34)
\end{aligned}$$

この推定方程式より $\hat{\pi}$ は、以下の式を満たすことがわかる。

$$\pi(1-\pi)(1-2\pi)^4(W-3Y^2)^3 = (1-6\pi+6\pi^2)^3 Z^4 \tag{4.3.35}$$

このように、π の6次式となり、$\hat{\pi}$ を代数的に求めることはできない。また、前節と同様に identifiability の欠如のため (4.3.35) は $\pi=\frac{1}{2}$ で対称となっている。よって、π を $\frac{1}{2}\le\pi\le 1$ に制限して以下議論を進める。

ここで、$H_0\ :\ \pi=1$ の下では、

$$\hat{\pi}(1-\hat{\pi})(1-2\hat{\pi})^4 = O_p(\frac{1}{\sqrt{n}})$$

である。推定方程式の解のうち最も1に近い値を与える解を $\hat{\pi}$ つまり、

$$\sqrt{n}(\hat{\pi}-1) = O_p(1)$$

と定義する。このことより、推定方程式 (4.3.35) より、

$$\begin{aligned}
\hat{\pi}-1 &= -\frac{Z^4(1+O_p(\frac{1}{\sqrt{n}}))}{(W-3Y^2)^3(1+O_p(\frac{1}{\sqrt{n}}))} \\
\sqrt{n}(\hat{\pi}-1) &= -\frac{(\sqrt{n}Z)^4}{\{\sqrt{n}(W-3Y^2)\}^3}+O_p(\frac{1}{\sqrt{n}}) && (4.3.36)
\end{aligned}$$

が得られる。H_0 の下での漸近分布は以下のようになる。

$$-\sqrt{n}(\hat{\pi}-1) \xrightarrow{\mathcal{L}} \frac{N^4(0,15\sigma^6)}{N^3(0,24\sigma^8)}$$

つまり、

$$-\frac{16\sqrt{6}}{75}\sqrt{n}(\hat{\pi}-1) \xrightarrow{\mathcal{L}} \frac{N^4(0,1)}{N^3(0,1)} \tag{4.3.37}$$

となる。

また、$0 \le \pi \le 1$ という制約をつけた場合の $\hat{\pi}$ の漸近分布は、前節と同様にして、

$$P\{-\frac{16\sqrt{6}}{75}\sqrt{n}(\hat{\pi}^*-1) < x\} \xrightarrow{\mathcal{P}} \frac{1}{2}+\frac{1}{2}P\{\frac{N^4(0,1)}{|N^3(0,1)|} < x\} \tag{4.3.38}$$

が得られる。

次に、この検定の理論的検出力を求めるため、この統計量 $\frac{16\sqrt{6}}{75}\sqrt{n}(\hat{\pi}-1)$ の H_1 の下での漸近分布を考える。まず、H_1 の下での (4.3.35) の解の挙動を調べるために $\pi=\pi_0$ のときの $\hat{\pi}$ の解を求めてみよう。このとき、

$$\begin{aligned} Z &\xrightarrow{\mathcal{P}} \pi_0(1-\pi_0)(1-2\pi_0)(\mu_1-\mu_2)^3 \\ W-3Y^2 &\xrightarrow{\mathcal{P}} \pi_0(1-\pi_0)(1-6\pi_0+6\pi_0^2)(\mu_2-\mu_1)^4 \end{aligned}$$

であるから、Z と $W-3Y^2$ をその収束先で置き換えると (4.3.35) は、

$$\pi(1-\pi)(1-2\pi)^4(1-6\pi_0+6\pi_0^2)^3-\pi_0(1-\pi_0)(1-2\pi_0)^4(1-6\pi+6\pi^2)^3=0$$

となり、これを解くと $\hat{\pi}$ の実数解は

$$\begin{aligned} \pi_0=\frac{1}{2} &\Rightarrow \hat{\pi}=0,\frac{1}{2}(\text{4 重根}),1 \\ \frac{1}{2}<\pi_0<\frac{2}{3} &\Rightarrow \hat{\pi}<0,\hat{\pi}=1-\pi_0,\pi_0,\hat{\pi}>1 \\ \frac{2}{3}\le\pi_0<1 &\Rightarrow \hat{\pi}=1-\pi_0,\pi_0 \end{aligned}$$

となることがわかる。

$\frac{2}{3}\le\pi_0<1$ のとき (4.3.35) の実数解は、$\hat{\pi}$ を 1 に最も近い解を選ぶという規準はうまく働くようである。つまり、$\hat{\pi} \xrightarrow{\mathcal{P}} \pi_0$ は成立すると思われる。

$\frac{1}{2}\le\pi_0<\frac{2}{3}$ のときは問題がある。(4.3.35) は $\pi=\pi_0, 1-\pi_0$ の近傍と、0 以下、1 以上にそれぞれ 1 つで計 4 つの解が得られ、$\hat{\pi}$ を 1 に最も近い解を選ぶという規準はうまく働かない。この場合はさらに詳しい検討が必要である。以下、$\frac{2}{3}\le\pi_0<1$ とし、$\hat{\pi}$ の一致性が成立すると仮定して議論を進める。

統計量 $-\frac{16\sqrt{6}}{75}\sqrt{n}(\hat{\pi}-1)$ の H_1 の下での漸近分布を考える。

$$-\frac{16\sqrt{6}}{75}\sqrt{n}(\hat{\pi}-\pi) \xrightarrow{\mathcal{L}} N(0,\frac{1536}{5625}\sigma_{\hat{\pi}}^2) \tag{4.3.39}$$

が得られる。ここで、$d=(\mu_2-\mu_1)/\sigma$ とおくと、$\sqrt{n}(\hat{\pi}-\pi)$ の漸近分散は、

$$\begin{aligned}\sigma^2_{\hat{\pi}} &= \pi(1-\pi)(1-24\pi^2+48\pi^3-24\pi^4)^2 \\ &+\frac{24\pi(1-\pi)-144\pi^2(1-\pi)^2-576\pi^3(1-\pi)^3+4032\pi^4(1-\pi)^4}{d^2} \\ &+\frac{216\pi(1-\pi)-2304\pi^2(1-\pi)^{+}8640\pi^3(1-\pi)^3}{d^4} \\ &+\frac{240-2016\pi(1-\pi)+5184\pi^2(1-\pi)^2}{d^6} \\ &+\frac{216-864\pi(1-\pi)}{d^8}+o(1)\end{aligned}$$

と表される。以上より $-\frac{16\sqrt{6}}{75}\sqrt{n}(\hat{\pi}-1)$ の H_1 の下での分布は、

$$N(-\frac{16\sqrt{6}}{75}\sqrt{n}(\pi-1),\frac{1536}{5625}\sigma^2_{\hat{\pi}}) \tag{4.3.40}$$

で近似される。

5　検出力の数値例

ここでは、上記のモーメント法による危険率 α の各検定の検出力をコンピュータシュミレーションにより求め、理論値との比較を行う。標本の大きさを n、有意水準を $\alpha=0.1$ と設定し、くり返し数 $r=100$ の数値実験に基づき、以下の各 $\pi, d=(\mu_2-\mu_1)/\sigma$ の組み合わせに対する検出力を求めた。

（１）$\mu_1(=0),\sigma^2(=1)$ が既知の場合。

理論値

$n=100$, π

d	0.2	0.5	0.8
0.5	0.75	0.61	0.48
1.0	1.00	0.84	0.44
1.5	1.00	1.00	0.41

$n=400$, π

d	0.2	0.5	0.8
0.5	0.95	0.82	0.59
1.0	1.00	1.00	0.75
1.5	1.00	1.00	0.87

シュミレーション（H_0 の下、100 回の中 10 回 H_0 が棄却）

$n=100$, π

d	0.2	0.5	0.8
0.5	0.85	0.55	0.25
1.0	0.99	0.96	0.39
1.5	1.00	1.00	0.40

$n=400$, π

d	0.2	0.5	0.8
0.5	1.00	0.94	0.48
1.0	1.00	1.00	0.76
1.5	1.00	1.00	0.88

（２）$\sigma^2(=1)$ が既知の場合。

理論値

$n = 100$, π

d	0.5	0.7	0.8
0.5	0.52	0.50	0.50
1.0	0.62	0.54	0.50
1.5	0.78	0.59	0.51

$n = 400$, π

d	0.5	0.7	0.8
0.5	0.54	0.52	0.51
1.0	0.78	0.66	0.58
1.5	0.98	0.86	0.72

シュミレーション（H_0 の下、100 回の中 11 回 H_0 が棄却）

$n = 100$, π

d	0.5	0.7	0.8
0.5	0.31	0.26	0.22
1.0	0.48	0.35	0.27
1.5	0.84	0.53	0.42

$n = 400$, π

d	0.5	0.7	0.8
0.5	0.36	0.25	0.27
1.0	0.94	0.74	0.61
1.5	1.00	0.97	0.94

このように、（１），（２）では、π が小さく $d = (\mu_2 - \mu_1)/\sigma$ が大きくなれば検出力は増加していき、また事前情報が少なくなるにつれ検出力は低下していくことがわかる。

（３）適合度検定

既存の検定として、適合度検定（χ^2 検定）との比較を行う。（$\sigma^2(=1)$ が既知、分割数が 8 とした）

シュミレーション（H_0 の下、100 回の中 10 回 H_0 が棄却）

$n = 100$, π

d	0.5	0.7	0.8
0.5	0.09	0.11	0.14
1.0	0.19	0.17	0.17
1.5	0.65	0.45	0.24

この方法と比較すると、モーメント法による検定の方が検出力は高いようである。

6 参考文献

Aitkin, M. and Rubin, D.B.(1985). Estimation and hypothesis testing in finite mixture models. *J.Royal Statist. Soc.*, **47**, 67-75.

Chen, J. and Kalbfleisch, J.D.(1993). Penalized minimum distance estimates in finite mixture models. *Technical Report Series*, **94-08** , University of Waterloo.

Chernoff, H.(1954). On the distribution of the likelihood ratio. *Ann. Math. Statist.*, **25**, 573-578.

Hartigan, J.A.(1985). A failure of likelihood asymptotics for normal mixture. In Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer, Vol. 2. edited by L. LeCam and R.A. Olshen, 807-810.

Henna, J.(1985). On estimating of the number of constituents of a finite mixture of continuous distributions. *Ann. Inst. Statist. Math.*, **37**, 235-240.

Leroux, B.(1992). Consistent estimation of a mixture distribution. *Ann. Statist.*, **20**, 1350-1360.

Shapiro, A.(1985). Asymptotic distribution of test statistics in the analysis of moment structures under inequality constraints. *Biometrika*, **72**, 133-144.

Shapiro, A.(1988). Towards a unified theory of inequality constrained testing in multivariate analysis. *Inter. Statist. Rev.*, **56**, 49-62.

Titterington, D.M., Smith, A.F.M. and Makov, U.E.(1985). *Statistical Analysis of Finite Mixture Distributions*. Wiley: New York.