# 統計的仮説検定

大橋耕也

東京工業大学
情報理工学院　数理・計算科学専攻

2017/10/23

# 統計的仮説検定とは？

- 統計的仮説検定とは、
  - 母集団分布の母数に関する仮説を標本から検証する統計学的方法のひとつ [引用: wikipedia]
  - つまり、「サンプルをもとに、ある仮説 (帰無仮説 (N.H.)) の正しさを、対立仮説 (A.H.) との比較により検証する手法」
  - 通常は、帰無仮説を棄却することを期待して行われる。
    （例　N.H.：新薬に有効性はない v.s. A.H.：新薬に有効性はある）
    （　　N.H.：データに不正はない v.s. A.H.：データに不正がある）

しかし、p-value abuser が多すぎた。。。

- p 値の使用を禁止するジャーナルも
  Psychology journal bans P values.2015 [7]
- 有意水準を.05 から.005 にしよう！との動きも

  **One Sentence Summary:** We propose to change the default *P*-value threshold for statistical significance for claims of new discoveries from 0.05 to 0.005.

  [8]

殺傷事件の責任は包丁にはありません (´・ω・`)

# 統計的仮説検定の枠組み

　サンプル $X^n$ に対して、ある検定統計量 $S(X^n)$ を定義し、ある定数 $a$ を用いて

$$
\begin{aligned}
S(X^n) \leq a &\Rightarrow select\ N.H. \\
S(X^n) > a &\Rightarrow reject\ N.H.
\end{aligned}
$$

とすることを「検定をする」と言います。$S(X^n) > a$ は、「サンプル $x^n$ が帰無仮説から発生している (真の分布が帰無仮説) と考えるにしては無理がある」という気持ちで、帰無仮説を棄却しています。
　したがって、検定統計量 $S(X^n)$ と定数 $a$ を決めることを「検定をつくる」と言い、様々な検定が提案されています。

　では、良い検定と悪い検定はどういった評価指数にもとづいて決められるのでしょうか？

# 良い検定と悪い検定

２つの量を導入します。

- 有意水準 (レベル)、または p 値

$$Level(S,a) = P(S(X^n) > a|N.H.)$$

N.H. が正しいにもかかわらず、棄却してしまう確率

- 検出力 (パワー)

$$Power(S,a) = P(S(X^n) > a|A.H.)$$

A.H. が正しいときに、ちゃんと棄却してくれる確率

これらの量は検定、すなわち $(S,a)$ を設計すれば定まる量です。
検定 $(S_1,a_1)$, $(S_2,a_2)$ に対して

$$Level(S_1,a_1) = Level(S_2,a_2) \Rightarrow Power(S_1,a_1) \geq Power(S_2,a_2)$$

が成り立つときに、$S_2$ よりも $S_1$ は良い検定である (強力である) といい、任意の検定よりも強力な検定を最強力検定といいます。

# 最強力検定

N.H. と A.H. が確率分布の形で与えられる場合、$S =$(分配関数比) が最強力検定であることが証明できます。

## 仮説が確率分布の形で与えられるとは？

確率分布の形で与えられる例

$$N.H. : p(x|w_0) \text{ v.s. } A.H. : p(x|w_1)$$

確率分布の形で与えられない例 (複合仮説)

$$N.H. : p(x|w_0) \text{ v.s. } A.H. : p(x|w \neq w_0)$$

なお、仮説が確率分布の形で与えられない場合は一般に最強力検定は存在しません。
2 つの事前分布に対して検定を行う、ベイズ検定を導入します。

# ベイズ検定

事前分布に対する仮説検定をベイズ検定という。
確率モデル $p(x|w)$ の事前分布に対して仮説を設定する。

$$N.H. : \varphi_0(w) \text{ v.s. } A.H. : \varphi_1(w)$$

みなさんお馴染みの形には

$$N.H. : \varphi_0(w) = \delta(w - w_0) \text{ v.s. } A.H. : \varphi_1(w) = \delta(w - w_1)$$

のように、デルタ関数を用いれば良いです。

では、分配関数の比が最強力検定を与えることを証明します。

# 最強力検定

**定理**

検定統計量

$$L(X^n) = \frac{\int dw \varphi_1(w) \prod_{i=1}^n p(X_i|w)}{\int dw \varphi_0(w) \prod_{i=1}^n p(X_i|w)}$$

は、最強力検定を与える。

(証明) $(L, b)$ と任意の検定 $(S, a)$ を考える．ただし，定数 $a, b$ は

$$P(L(X^n) > b|\varphi_0) = P(S(X^n) > a|\varphi_0)$$

を満たすものとする．このとき，次の不等式を示せばよい：

$$P^* \equiv P(L(X^n) > b|\varphi_1) - P(S(X^n) > a|\varphi_1) \geq 0.$$

ここで，それぞれの検定の棄却域 $A, B$ を定義する．

$$A = \{x^n;\ S(x^n) > a\}$$
$$B = \{x^n;\ L(x^n) > b\}$$

# 最強力検定

このとき，

$$P^* = \int \varphi_1(w)dw \left[\int_B - \int_A\right] \prod_{i=1}^n p(x_i|w)dx_i$$

$$= \int \varphi_1(w)dw \left[\left(\int_B - \int_{A\cap B}\right) - \left(\int_A - \int_{A\cap B}\right)\right] \prod_{i=1}^n p(x_i|w)dx_i$$

$$= \int \varphi_1(w)dw \left[\int_{B\setminus A} - \int_{A\setminus B}\right] \prod_{i=1}^n p(x_i|w)dx_i$$

集合 $B \setminus A$ において，$L(x^n) > b$ であるから

$$\int \varphi_1(w)dw \int_{B\setminus A} \prod_{i=1}^n p(x_i|w)dx_i > b \int \varphi_0(w) \int_{B\setminus A} \prod_{i=1}^n p(x_i|w)dx_i$$

同様に，集合 $A \setminus B$ においては

$$\int \varphi_1(w)dw \int_{A\setminus B} \prod_{i=1}^n p(x_i|w)dx_i \leq b \int \varphi_0(w) \int_{A\setminus B} \prod_{i=1}^n p(x_i|w)dx_i$$

# 最強力検定

以上から，

$$
\begin{aligned}
P^* &\geq b \int \varphi_0(w)dw \left[ \int_{B \setminus A} - \int_{A \setminus B} \right] \prod_{i=1}^{n} p(x_i|w)dx_i \\
&= b \int \varphi_0(w)dw \left[ \int_{B} - \int_{A} \right] \prod_{i=1}^{n} p(x_i|w)dx_i \\
&= 0
\end{aligned}
$$

と不等式が示されたので，検定 $(L, b)$ は最強力検定である．(証明終わり)

特に帰無仮説，対立仮説ともに事前分布としてデルタ関数を用いた場合に，この定理はネイマン・ピアソンの補題と呼ばれます．
定理から，我々は分配関数の比 (周辺尤度比) の確率分布を導出できれば，最強力検定をつくれることになります．

# 正則モデルに対する検定

**定理**

確率モデル $p(x|w)$ $(w \in \mathbb{R}^d)$ が正則であるとき，仮説

$$
\begin{aligned}
N.H. &: \varphi_0(w) = \delta(w - w_0) \\
A.H. &: \varphi_1(w)
\end{aligned}
$$

に対する対数周辺尤度比は自由度 $d$ のカイ二乗分布へ法則収束する．

正則モデルである場合，サンプル $X_1, X_2, \ldots, X_n$ を生成するモデルのパラメータが $w_0$ であるか否かの最強力検定はカイ二乗検定となることを示している．

# 正則モデルに対する検定

（証明）
確率モデルが正則な場合，経験誤差関数は以下で与えられる [1]

$$\frac{1}{n}\sum_{i=1}^{n}\log\frac{p^*(X_i)}{p(X_i|w_0)} = \frac{1}{n}\left(\frac{1}{2}\|\xi_n\|^2 + o_p(1)\right)$$

**注意**

検定とは,「真の分布が N.H. である世界から見たときに今手元にあるサンプルは偶然の偏りと考えるにしても無理があるなぁ...」という思想なので真の分布は確率モデルで実現可能である．つまり，$d = \mathrm{tr}(IJ^{-1})$

したがって，

$$L(X^n) = \log\frac{p^*(X^n)}{p(X^n|w_0)} \xrightarrow{d} \frac{1}{2}\chi_d^2$$

(証明終わり)

# 特異モデルに対する検定

正則モデルの場合には，対数周辺尤度比がカイ二乗分布へ法則収束するため検定をつくることができました．しかし，特異モデルの場合は確率項の分布は一般に不明です．[1]

$$F_n(1) = nL_n(w_0) + \lambda \log n - (m-1)\log\log n + \Theta_n(\beta, \xi_n) + o_p(1)$$

したがって，特異モデルに対して検定をつくる場合は上記の赤い項の確率分布を求める必要があります．

今のところは各問題に対して個々に確率分布を考える必要があり，私の研究テーマになります．[6]

# 有名な検定の紹介

以下のよく知られた（学部で習うであろう）検定手法を紹介します.

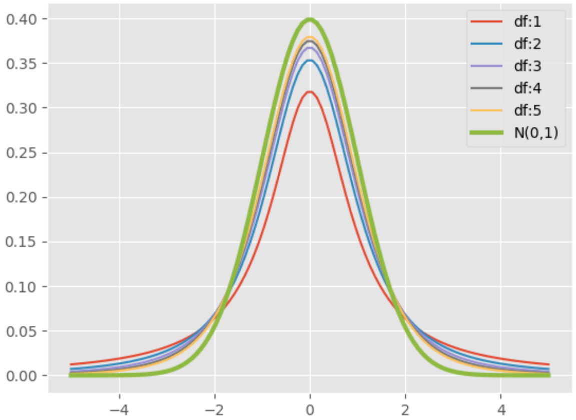
- 正規分布の平均の検定（z 検定）
- 正規分布の平均の検定（t 検定）
- カイ二乗適合度検定
- カイ二乗独立性検定
- 分散分析（ANOVA）

各検定手法で，データにどのような仮定を置いているのかに注意しましょう.

# 検定によく用いられる分布

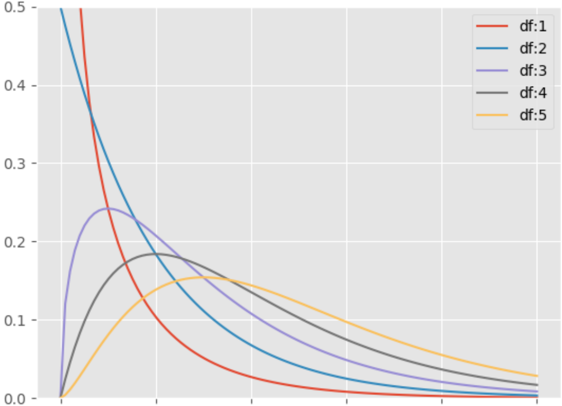
ここでは，t 分布，カイ二乗分布，エフ分布を定義します．

- t 分布
  自由度 $n$ の t 分布の p.d.f $t_n = N(0, 1^2)/\sqrt{\chi_n^2/n}$
  正規分布に似ていますが，裾が重い分布です．

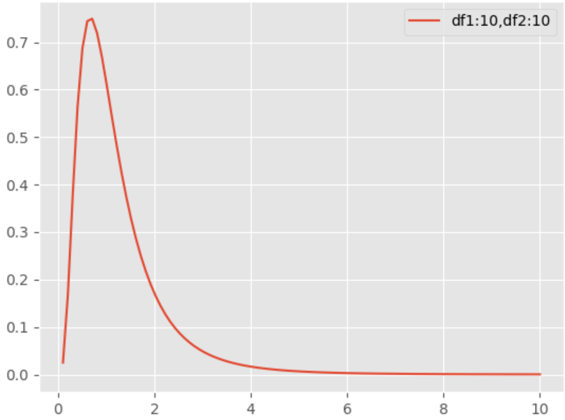


# 検定によく用いられる分布

- カイ二乗分布
  $Z_i \sim N(0, 1^2)$ に対して，確率変数
  $$Y = Z_1^2 + Z_2^2 + \cdots + Z_n^2$$
  の従う分布を自由度 $n$ のカイ二乗分布といいます.



- エフ分布
  自由度 $(m,n)$ のエフ分布の p.d.f $F_n^m = \left(\chi_m^2/m\right) / \left(\chi_n^2/n\right)$



# 正規分布の平均の検定（z 検定）

**仮定**

統計モデル $p(x|w) = N(x|\mu, \sigma^2)$，ただし分散はあらかじめ与える．

このとき，サンプル $X_1, \ldots, X_n$ をもとに次を検定したい．

$$
\begin{aligned}
N.H. &: \mu = \mu_0 \\
A.H. &: \mu = \mu_1 (> \mu_0)
\end{aligned}
$$

では，対数尤度比を計算してみましょう．

$$
\begin{aligned}
\log \frac{\prod p(X_i|\mu_1, \sigma^2)}{\prod p(X_i|\mu_0, \sigma^2)} &= \sum \log \left( \exp \left( -\frac{1}{2}(X_i - \mu_1)^2 + \frac{1}{2}(X_i - \mu_0)^2 \right) \right) \\
&= \frac{1}{2} \sum \left( \mu_0^2 - \mu_1^2 + 2(\mu_1 - \mu_0) X_i \right)
\end{aligned}
$$

確率的に揺れている項は $\bar{X} = \sum X_i / n$ のみなので，この量を使って検定がつくれます．

# 正規分布の平均の検定（z 検定）

今，N.H. を仮定しているので，サンプルは $N(\mu_0, \sigma^2)$ に従います．したがって，次のように標準化ができます．

$$Z \equiv \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1^2)$$

以上から，正規分布の平均に関する検定は $S(X^n) = Z$，定数 $a$ は標準正規分布の分布表から決定できる量でつくることができました．
これを z 検定といいます．

# 正規分布の平均の検定（t 検定）

**仮定**

統計モデル $p(x|w) = N(x|\mu, \sigma^2)$

z 検定において，分散も未知である場合に平均の検定を行いたい．

$$
\begin{aligned}
N.H. &: \mu = \mu_0 \\
A.H. &: \mu = \mu_1 (> \mu_0)
\end{aligned}
$$

このとき，分散の不偏推定量

$$
\hat{\sigma}^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2
$$

を用いて次のような変換を施します．

$$
T \equiv \frac{\bar{X} - \mu_0}{\hat{\sigma}/\sqrt{n}}
$$

# 正規分布の平均の検定（t 検定）

**定理**

$T$ は自由度 $n-1$ のティー分布に従う．

（証明）

$X_1, \dots, X_n \sim N(\mu, \sigma^2)$ とする．

$$T = \frac{\bar{X}-\mu}{\hat{\sigma}/\sqrt{n}} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \times \frac{\sigma}{\hat{\sigma}} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \times 1/\sqrt{\frac{\sum(X_i-\bar{X})^2}{\sigma^2} \times \frac{1}{n-1}}$$

t 分布の定義より，赤い項が $\chi^2_{n-1}$ に従うことを示せばよい．

# 正規分布の平均の検定（t 検定）

$$\begin{aligned}\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i^2-2X_i\mu+\mu^2)\\ &= \frac{1}{\sigma^2}\left(\sum_{i=1}^{n}X_i^2-2n\bar{X}\mu+n\mu^2\right)\\ &= \left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2+\frac{\sum X_i^2-n\bar{X}^2}{\sigma^2}\\ &= \left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2+\frac{\sum(X_i-\bar{X})^2}{\sigma^2}\end{aligned}$$

したがって，標準正規分布に従う確率変数を $Z_i$ と表記すると

$$\frac{\sum(X_i-\bar{X})^2}{\sigma^2}=\sum_{i=1}^{n-1}Z_i^2\sim\chi^2_{n-1}$$

が示された.（証明終わり）

# 正規分布の平均の検定（t 検定）

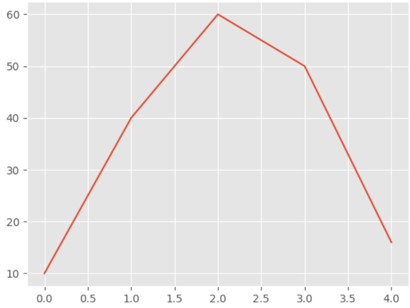
以上から，分散が未知である場合，t 変換

$$T \equiv \frac{\bar{X} - \mu_0}{\hat{\sigma}/\sqrt{n}}$$

を施し，この統計量が自由度 $n-1$ のティー分布に従うことを利用して検定を行うことができます．
これを t 検定といいます．

# カイ二乗適合度検定

| 回数 $x$ | 0 | 1 | 2 | 3 | 4 | 計 |
|---|---|---|---|---|---|---|
| 頻度 $f(x)$ | 10 | 40 | 60 | 50 | 16 | 176 |




このサンプルから，$f(x)$ は二項分布 $B(4, p)$ となるかどうかを検定したい．

# カイ二乗適合度検定

サンプル $X_1, \ldots, X_n$ が $f(x|\theta)$ $(\theta \in \mathbb{R}^s)$ に従うかどうかを検定したい．
互いに素な事象 $A_1, \ldots, A_k$ に対して，確率分布

$$\pi(\theta) = (\pi_1(\theta), \ldots, \pi_k(\theta))$$

$$\pi_i(\theta) = \int_{A_i} f(x|\theta)dx$$

を考えます．また，パラメータには最尤推定量 $\hat{\theta}$ を用います．
このとき，$n$ 個のサンプルが $f(x|\hat{\theta})$ に従うならば，事象 $A_i$ は $n\hat{\pi}_i$ 回起こると期待されます．

| 事象 | $A_1$ | $\cdots$ | $A_k$ |
|---|---|---|---|
| 実現値 | $n_1$ | $\cdots$ | $n_k$ |
| 期待値 | $n\hat{\pi}_1$ | $\cdots$ | $n\hat{\pi}_k$ |

# カイ二乗適合度検定

| 事象 | $A_1$ | $\cdots$ | $A_k$ |
|---|---|---|---|
| 実現値 | $n_1$ | $\cdots$ | $n_k$ |
| 期待値 | $n\hat{\pi}_1$ | $\cdots$ | $n\hat{\pi}_k$ |

このとき，次の検定統計量を考えます．

$$S(X^n) = \sum_{i=1}^{k} \frac{(n_i - n\hat{\pi}_i)^2}{n\hat{\pi}_i}$$

つまり，どれだけサンプルが期待値に対してずれているかを表す統計量であり次の定理が成り立ちます．

**定理**

$$S \to \chi^2_{(k-1-s)} \text{ as } n \to \infty$$

# カイ二乗適合度検定

では，最初の例を適合度検定してみましょう．

| 回数 $x$ | 0 | 1 | 2 | 3 | 4 | 計 |
|---|---|---|---|---|---|---|
| 頻度 $f(x)$ | 10 | 40 | 60 | 50 | 16 | 176 |

二項分布の最尤推定量は

$$\hat{p} = \frac{\bar{x}}{4} \fallingdotseq 0.53$$

であり，

$$\hat{\pi}_i = B(i|4, 0.53)$$

であるので，それぞれの事象の期待される生起数は

$$n\hat{\pi} = (8.6, 38.7, 65.5, 49.3, 13.9)$$

となります．検定統計量を自由度 3 のカイ二乗分布の有意水準 0.05 点により検定しましょう．

$$S(x^{176}) \fallingdotseq 1.07 < \chi^2_{3,0.05} \fallingdotseq 7.81$$

したがって，帰無仮説は棄却されずサンプルは二項分布に従わないとは言えません

# カイ二乗適合度検定

疑問：

　検定は帰無仮説の採択にはあまり意味をなさないのに，これをもって二項分布に従うと結論付けて良いのでしょうか？　分布の適合度検定には，コルモゴロフスミルノフ検定と呼ばれるものもありますが，採択をもって結論付ける点では同じです.

２標本コルモゴロフ-スミルノフ検定は，２群データの従う分布が等しいかどうか（棄却をもって，２郡の分布は違うといえる）を検定するときに使われるみたいです.

# カイ二乗独立性検定

カイ二乗適合度検定を応用すると，各事象の独立性検定を行うことができます.

$$N.H. : p(A_i \cap B_j) = p(A_i)p(B_j)$$
$$A.H. : otherwise$$

たとえば，A=$\{$ 喫煙，非喫煙 $\}$，B=$\{$ 飲酒，非飲酒 $\}$ としたとき，それらが独立かどうかをカイ二乗検定により検定できます.

# 分散分析（ANOVA）

３群以上からなる母平均の差を検定したい場合，分散分析（analysis of variance）が用いられます．例えば，都市の大きさ（５００万都市，１００万都市，５０万都市など）により小学生の学力に差があるのか？といった問いに答えられます．
都市の大きさを要因といい，$r$ 水準に分けられたとしましょう．

$$
\begin{aligned}
A_1 :\ & Y_1 = (Y_{11}, \ldots, Y_{1n_1}) \sim N(\mu_1, \sigma^2) \\
A_2 :\ & Y_2 = (Y_{21}, \ldots, Y_{2n_2}) \sim N(\mu_2, \sigma^2) \\
& \vdots \\
A_r :\ & Y_r = (Y_{r1}, \ldots, Y_{rn_r}) \sim N(\mu_r, \sigma^2)
\end{aligned}
$$

このとき，分散分析では次の仮説を置きます．

$$
\begin{aligned}
& N.H. : \mu_1 = \mu_2 = \cdots = \mu_r \\
& A.H. : otherwise
\end{aligned}
$$

# 分散分析（ANOVA）

**定理**

$$F = \frac{\frac{\sum n_i(\bar{Y}_i - \bar{Y})^2}{r-1}}{\frac{\sum\sum(Y_{ij} - \bar{Y}_i)^2}{n-r}} \sim F^{r-1}_{n-r}$$

- この検定統計量は，水準内での分散に対して水準間での分散がどの程度大きいのかを表しており，エフ分布を用いて検定ができます．
- 各水準間での分散が相対的に大きいならば，水準によりデータは影響を受け帰無仮説は棄却されるべきというのは感覚的にも理解しやすいですね．
- サンプルサイズを大きくすれば，$F$ は大きくなり帰無仮説は棄却できることに注意です．これは，群間に有意差があるかどうか？はサンプルを頑張って集めてくれば必ず示せることを示唆しており，注意すべき点です．　50 人では有意差は確認されなかったのに，100 人では有意差が確認されることも普通にあります．それを理解して使いましょう．

# 特異モデルに対する検定

以上，よく知られた検定を紹介してきましたが基本的に統計モデルとして正則モデルを用いており，よく知られた分布（標準正規分布，t 分布，カイ二乗分布，エフ分布等）により検定をつくることができていました．

ここで，特異モデルに対するベイズ検定を変化点検出へ応用した論文 [藤原，渡辺 2008] を紹介します．

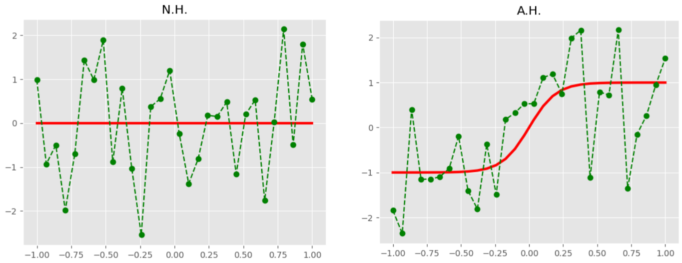
# ベイズ検定による変化点検出 [6]

## 変化点検出

統計モデル: $y_i = a\tanh(bx_i) + N(0,\sigma^2)$
ただし，$x_1,\dots,x_n$ は固定点とする．これに対し次の仮説を設定する．

$$N.H.: \delta(a)\delta(b)$$

$$A.H.: \begin{cases} 1/4 & (|a|\le 1, |b|\le 1, ab\ne 1) \\ 0 & (otherwise) \end{cases}$$



# ベイズ検定による変化点検出 [6]

この統計モデルは，パラメータ $w=\{a,b\}$ からモデルへの写像が単射ではないため，特異モデルです．まずは，周辺尤度比を計算しましょう．

$$
\begin{aligned}
L(Y^n) &= \frac{\int e^{-\frac{1}{2\sigma^2}\sum(y_i - a\tanh(bx_i))^2}\frac{1}{4}dw}{e^{-\frac{1}{2\sigma^2}\sum y_i^2}} \\
&= \int \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} a^2\tanh(bx_i)^2 - 2\sum_{i=1}^{n} y_i a\tanh(bx_i)\right)\right)\frac{1}{4}dadb
\end{aligned}
$$

サンプルは帰無仮説に従うと仮定しているので，サンプルサイズが大きくなるにつれて事後分布は特異点である原点 $(a,b)=(0,0)$ の近傍に集中していきます．原点の近傍でテイラー展開すると

$$
a\tanh(bx) = abx + c_1ab^3x^3 + c_2ab^5x^5 + \cdots
$$

となります．

# ベイズ検定による変化点検出 [6]

$$a \tanh(bx) = abx + c_1 ab^3 x^3 + c_2 ab^5 x^5 + \cdots$$

第二項の係数 $\{ab^3, ab^5, \ldots\}$ は第一項の係数 $ab$ のイデアルに含まれるため，漸近分布には影響しません．つまり，$a \tanh(bx) \simeq abx$ として $L(Y^n)$ の漸近分布を議論できることになります．

$$L(Y^n) \simeq \int e^{-nNa^2b^2 + \sqrt{n}Mab} \frac{1}{4} dadb$$

ただし，

$$N = \frac{1/n \sum x_i^2}{2\sigma^2} \quad : \quad const.$$
$$M = \frac{1/\sqrt{n} \sum x_i y_i}{\sigma^2} \sim N(0, 2N)$$

です．

# ベイズ検定による変化点検出 [6]

デルタ関数を用いて式変形をしていきます.

$$
\begin{aligned}
L(Y^n) &= \int_{-1}^{1} dt \int_{[-1,1]^2} \frac{dadb}{4} \delta(t-ab) e^{-nNt^2+\sqrt{n}Mt} \\
&= \int_{-\sqrt{n}}^{\sqrt{n}} \frac{dt}{\sqrt{n}} \int_{-1}^{1} da \int_{-1}^{1} db \frac{1}{4} \delta\left(\frac{t}{\sqrt{n}} - ab\right) e^{-Nt^2+Mt}
\end{aligned}
$$

## 定理

$$
\int_{-1}^{1} da \int_{-1}^{1} db \left(\frac{t}{\sqrt{n}} - ab\right) = -2\log\frac{|t|}{\sqrt{n}}
$$

この定理を使ってさらに計算を進めていきます.

# ベイズ検定による変化点検出 [6]

$$
\begin{aligned}
L(Y^n) &= \int_{-\sqrt{n}}^{\sqrt{n}} \frac{dt}{\sqrt{n}} \frac{1}{4} \left( -2\log \frac{|t|}{\sqrt{n}} \right) e^{-Nt^2+Mt} \\
&= \int_0^{\sqrt{n}} \frac{dt}{\sqrt{n}} \left( -\log \frac{|t|}{\sqrt{n}} \right) e^{-Nt^2} \left( \frac{e^{Mt}+e^{-Mt}}{2} \right) \\
&= \int_0^{\sqrt{n}} \frac{dt}{\sqrt{n}} \left( -\log \frac{|t|}{\sqrt{n}} \right) e^{-Nt^2} \cosh(Mt)
\end{aligned}
$$

これで，周辺尤度比がよく知られた分布に従う確率変数の関数で計算できることになります．

$M \sim N(0, 2N)$ なので，有意水準 0.05 の棄却域は平均 $\pm$ 標準偏差 $\times 2$ の外側で決定できます．$L(Y^n)$ は $M$ の偶関数かつ $|M|$ の単調増加関数であるから，

$$
\begin{aligned}
W &= \{|M| \geq 2\sqrt{2N}\} \\
&= \{-\log L(|M|) \leq -\log L(2\sqrt{2N})\} \\
&= \{-\log L(M) \leq -\log L(2\sqrt{2N})\}
\end{aligned}
$$

# ベイズ検定による変化点検出 [6]

以上の議論から，自由エネルギーの差分

$$-\log L(Y^n) = \frac{\int e^{-\frac{1}{2\sigma^2}\sum(y_i - a\tanh(bx_i))^2}\frac{1}{4}dw}{e^{-\frac{1}{2\sigma^2}\sum y_i^2}}$$

の値が，

$$-\log L(2\sqrt{2N}) = \frac{1}{2}\log n - \log\int_0^{\sqrt{n}}\left(-\log\frac{|t|}{\sqrt{n}}\right)e^{-Nt^2}\cosh(2\sqrt{2N}t)dt$$

以下であれば，有意水準 0.05 で帰無仮説を棄却することで検定が可能になりました．
ちなみに，この議論は一般の解析関数 $af(bx)$ に対しても適用できます．
変化点検出を応用先に設定した場合は tanh を用いるのが自然であると考えたのだと思います．

# 参考文献 I


[1] 渡辺澄夫. ベイズ統計の理論と方法. コロナ社. 2012

[2] 稲垣宣生. 数理統計学. 裳華房. 1990

[3] 福水健次, 栗木哲, 竹内啓, 赤平昌文. 特異モデルの統計学. 岩波書店. 2004

[4] 下平英寿, 伊藤秀一, 久保川達也, 竹内啓. モデル選択. 岩波書店. 2004

[5] 小寺平治. 明解演習 数理統計. 共立出版.1986

[6] 藤原香織; 渡辺澄夫. 特異モデルにおけるベイズ検定と時系列解析への応用. 電子情報通信学会論文誌 D, 2008, 91.4: 889-896.

[7] http://www.nature.com/news/psychology-journal-bans-p-values-1.17001

[8] Benjamin, Daniel J et al. “ Redefine Statistical Significance ”. PsyArXiv, 22 July 2017. Web.