母分散に関する検定

～Rを利用した検出力の比較～

20616028 宇佐美 友梨
20616014 石川 沙也夏
20616075 庄條 友子

概要

平均が0、分散がσ^2（未知の値）の正規母集団$N(0, \sigma^2)$からの大きさnの無作為標本$X_1, X_2, \ldots, X_n$があるとする。

n個の数値は、互いに独立である。

両側検定問題

$$H_0 : \sigma^2 = 1 \quad \text{vs.} \quad H_1 : \sigma^2 \neq 1$$

における尤度比検定 φ を定める。

H_1 においては、$\sigma^2 = \sigma_0^2$（ただし $\sigma_0^2 \neq 1$）として計算する。

さらに、2つの検出力をグラフで比較し、尤度比検定 φ が一様最強力検定であるかを調べる。

発表内容

- 帰無仮説 … とりあえず「真」と想定する仮説。
  H_0で表す。
- 対立仮説 … 帰無仮説を棄却するとき、代わりに採択する仮説。H_1で表す。

- 単純仮説 … ただ1点の母数(ただ1つの分布)を許容する仮説。
- 複合仮説 … 複数の分布を許容する仮説。

誤り：{ 第1種の誤り（帰無仮説が正しいのに棄却してしまう）
第2種の誤り（帰無仮説が誤っているのに採択してしまう）

		統計家の判定	
		H_0	H_1
真の仮説	H_0	正しい判定	第1種の誤り
	H_1	第2種の誤り	正しい判定

検定 φ が

$$E_{\sigma^2=1}[\varphi(X_1, X_2, \ldots, X_n)] \leq \alpha, \quad (0<\alpha<1)$$

をみたすなら φ は有意水準 α の検定という.

$\varphi(X_1, X_2, \ldots, X_n)$ の期待値をすべての σ^2 において考えた関数

$$\beta(\sigma^2; \varphi) = E_{\sigma^2}[\varphi(X_1, X_2, \ldots, X_n)]$$

$$= \int \varphi(x_1, x_2, \ldots, x_n) \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\sum_{i=1}^{n} \frac{x_i^2}{2\sigma^2}\right\} dx_1 dx_2 \ldots dx_n$$

を φ の検出力という.

一様最強力検定

$$E_{\sigma^2}[\varphi(X_1, X_2, \ldots, X_n)] \leq \alpha$$

をみたす検定φが水準αの任意の検定ψに対して

$$E_{\sigma^2}[\varphi(X_1, X_2, \ldots, X_n)] \geq E_{\sigma^2}[\psi(X_1, X_2, \ldots, X_n)]$$

をみたすならば、φを水準αの一様最強力検定という。

ネイマン・ピアソンの基本定理

$X_1, X_2, \ldots, X_n$ は正規母集団 $N(0, \sigma^2)$ からの大きさ n の無作為標本、その標本値を $x_1, x_2, \ldots, x_n$ とするとき、

検定問題

$$H_0 : \sigma^2 = 1 \quad \text{vs} \quad H_1 : \sigma^2 = \sigma_0^2$$

に対する水準 α の最強力検定 φ は

次式で与えられる.

$$\varphi(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & (L(1) < kL(\sigma_0^2)) \\ 0 & (L(1) > kL(\sigma_0^2)) \end{cases}$$

ここで、$L(1)$, $L(\sigma_0^2)$ は

$$L(1) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{\sum_{i=1}^{n} x_i^2}{2}\right)$$

$$L(\sigma_0^2) = \left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right)^n \exp\left(-\frac{\sum_{i=1}^{n} x_i^2}{2\sigma_0^2}\right)$$

となる尤度関数であり、定数 $k>0$ は

$$E_{\sigma^2=1}[\varphi(X_1, X_2, \ldots, X_n)] = \alpha$$

から定まるものである.

尤度比検定

$$\Lambda(x_1, x_2, \ldots, x_n) = \frac{L(1)}{\max_{\sigma^2} L(\sigma^2)} < \lambda$$

ならばH。を棄却し、そうでないなら受容。λは

$$P_{\sigma^2=1}(\Lambda(X_1, X_2, \ldots, X_n) < \lambda) = \alpha$$

をみたすように選ぶ。この検定方式を水準αの尤度比検定という。

$\Lambda = \Lambda(x_1, x_2, \ldots, x_n)$ … 尤度比

$X_1, X_2, \ldots, X_n$ は正規母集団 $N(0, \sigma^2)$ から大きさ n の無作為標本。

両側検定問題：

$$H_0 : \sigma^2 = 1 \quad \text{vs.} \quad H_1 : \sigma^2 \neq 1$$

H_1 において $\sigma^2 = \sigma_0^2$（ただし $\sigma_0^2 \neq 1$）とおく。

尤度比：$$\Lambda = \frac{L(1)}{\max_{\sigma_0^2 \neq 1} L(\sigma_0^2)} = \left(\frac{\ell}{n}\right)^{\frac{n}{2}} \exp\left\{-\frac{\ell - n}{2}\right\} < \lambda$$

$$\left(\ell = \sum_{i=1}^{n} x_i^2\right)$$

$$\Longleftrightarrow$$

$$\ell < c_1 \quad \text{または} \quad \ell > c_2$$

ただし、$c_1^{\frac{n}{2}} e^{-\frac{c_1}{2}} = c_2^{\frac{n}{2}} e^{-\frac{c_2}{2}}$ … ① をみたす。
水準αの棄却域をつくるには、

$$P_{\sigma^2=1}(S<c_1)+P_{\sigma^2=1}(S>c_2)=\alpha \quad \cdots ②$$

をみたすようにする。②をみたすようにc_1, c_2を定めるのは難しい。しかし、αとnに値を代入し、①と②の連立方程式を解けばc_1, c_2は決まる。

尤度比検定φの棄却域：

$$C=\{(x_1,x_2,\dots,x_n): s<c_1 \text{ または } s>c_2\}$$

$P_{\sigma^2=1}(S<c_1')=\frac{\alpha}{2}$, $P_{\sigma^2=1}(S>c_2')=\frac{\alpha}{2}$とおく場合の検定

$$P_{\sigma^2=1}(S<c_1')+P_{\sigma^2=1}(S>c_2')=\frac{\alpha}{2}+\frac{\alpha}{2}=\alpha$$

となるので水準αの棄却域をつくることができる。

検定φ'の棄却域：

$$C' = \{(x_1, x_2, \ldots, x_n) ; \ell < c_1' \text{ または } \ell > c_2'\}$$

H_1のもとで検出力は尤度比検定φと検定φ'において、

$$\beta(\sigma_0^2, \varphi) = P_{\sigma^2=\sigma_0^2}\left(\frac{S}{\sigma_0^2} < \frac{c_1}{\sigma_0^2}\right) + P_{\sigma^2=\sigma_0^2}\left(\frac{S}{\sigma_0^2} > \frac{c_2}{\sigma_0^2}\right)$$

$$\beta(\sigma_0^2, \varphi') = P_{\sigma^2=\sigma_0^2}\left(\frac{S}{\sigma_0^2} < \frac{c_1'}{\sigma_0^2}\right) + P_{\sigma^2=\sigma_0^2}\left(\frac{S}{\sigma_0^2} > \frac{c_2'}{\sigma_0^2}\right)$$

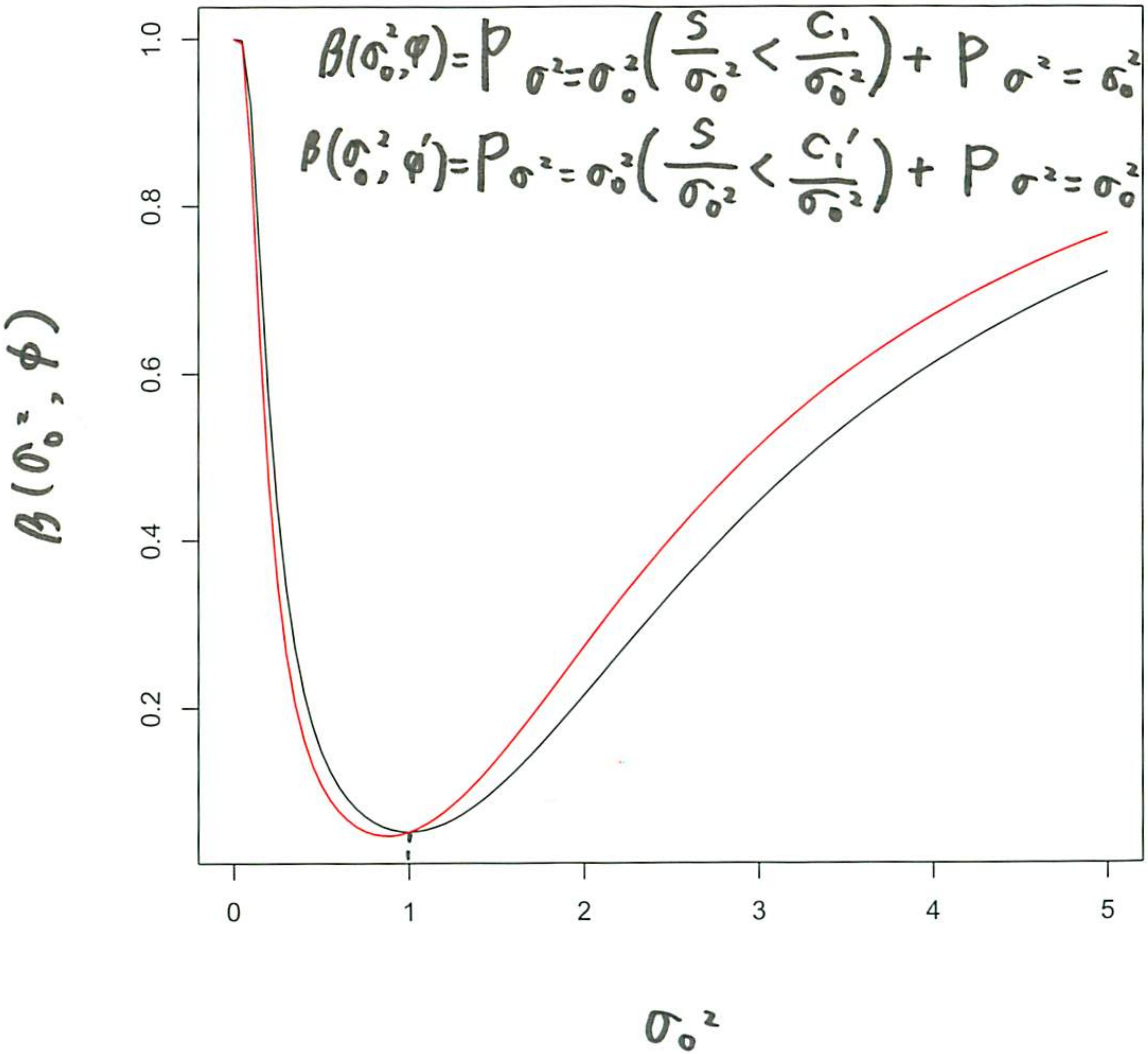
σ_0^2を横軸，$\beta(\sigma_0^2, \phi)$を縦軸としてグラフを表示。

$n=5,\ \alpha=0.05$ の検出力のグラフ　$C_1=0.9892$　$C_2=14.36879$

$C_1'=0.8312005$　$C_2'=12.83248$

$$\beta(\sigma_0^2,\varphi)=P_{\sigma^2=\sigma_0^2}\left(\frac{S}{\sigma_0^2}<\frac{C_1}{\sigma_0^2}\right)+P_{\sigma^2=\sigma_0^2}\left(\frac{S}{\sigma_0^2}>\frac{C_2}{\sigma_0^2}\right):$$

$$\beta(\sigma_0^2,\varphi')=P_{\sigma^2=\sigma_0^2}\left(\frac{S}{\sigma_0^2}<\frac{C_1'}{\sigma_0^2}\right)+P_{\sigma^2=\sigma_0^2}\left(\frac{S}{\sigma_0^2}>\frac{C_2'}{\sigma_0^2}\right):$$



まとめ

両側検定において、一様最強力検定は必ずしも存在するとは限らない、ということが分かった。

参照：
[連載]フリーソフトによるデータ解析・マイニング 第11回 Rと仮説検定
www1.doshisha.ac.jp/~mjin/R/11.pdf
検出力
aoki2.si.gunma-u.ac.jp/lecture/Kentei/power.html
[PDF] 1 検定論の枠組み
mcm-www.jwu.ac.jp/~konno/pdf/statha59.pdf
Inductive Statistics
http://www.wwq.jp/indsta/indsta24b.html
[PDF] 1 検定 2 検定の定式化
www.math.cm.is.nagoya-u.ac.jp/.../lec.../07.2007.11.29.test.pdf