推定統計量とその応用

統計 → 推定, 推論, 検定

検定:ある仮説を出して, その仮説の妥当性を確かめる

簡単な測定モデル

$$X_i = \theta + \varepsilon_i \quad (i = 1, 2, \cdots, N)$$

X_i :N ヶの測定結果　θ :未知パラメタ　ε_i :観測誤差(雑音)

(観測誤差)を確率変数として取り扱う.

$$\{\varepsilon_i\}\text{に関して}\begin{cases}\text{仮定1}: \varepsilon_1, \varepsilon_2, \cdots, \varepsilon_N \text{ が互いに独立} \\ \text{仮定2}: \text{同一分布} \\ \text{仮定3}: E\{\varepsilon_i\} = 0 \\ \text{仮定4}: E\{\varepsilon_i{}^2\} < \infty \quad \text{[分散が有界(有限値)]}\end{cases}$$

もし ε_i の確率分布が分かっていたら

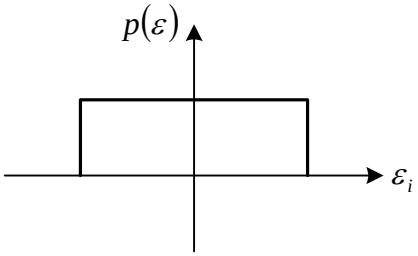
例:ピットマンの推定量

$$\hat{\theta} = \frac{1}{2}[\mathrm{Max}\ X_i + \mathrm{Min}\ X_i] \underset{\text{比較}}{\Leftrightarrow} \hat{\theta} = \frac{X_1 + \cdots + X_N}{N} \quad (\text{標本平均})$$

標本平均の分散は $O\left(\frac{1}{N}\right)$

O:order

もし



一様分布だったら, 推定分散は $O\left(\frac{1}{N^2}\right)$

なぜ確率分布が分かった時, ピットマン推定量がいいのかはこれから議論！

推定量と推定値

推定量(関数　$y = f(x)$)　推定値(関数値:具体的に x が与えられた時の y 値)

未知パラメタ θ の推定量

$$\underset{\text{確率変数}}{\hat{\theta}} = \hat{\theta}(X_1, \cdots, X_N)$$

$X_1, \cdots, X_N$:確率変数

$\hat{\theta}$確率的性質 $\begin{cases} E(\hat{\theta}) = \theta \rightarrow \text{不偏推定量（未知確定値）} \\ E\left[(\hat{\theta}-\theta)^2\right] \rightarrow \text{なるべく小さいことが望ましい} \end{cases}$　どこまで分散値を小さくすることができるか？

$$\underbrace{V(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right] \geq \frac{1}{NI}}_{\text{クラメール・ラオの不等式}}$$

$V(\hat{\theta})$:任意の推定量　$\hat{\theta}$の分散　N:データ数

但し，

$$I = E\left(\left(\frac{\partial}{\partial\theta}\log f(X,\theta_0)\right)^2\right) \quad \text{:フィッシャーの情報量}$$

$$= -E\left(\frac{\partial^2}{\partial\theta^2}\log f(X,\theta_0)\right)$$

（母数θを含んだ）確率変数 X の確率密度関数; $f(X,\theta)$

θ_0 :真の母数（パラメタ）

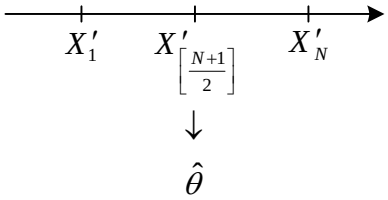
推定量の例　$X_i = \theta + \varepsilon_i \quad (i = 1,\cdots,N)$

1）標本平均

$$\hat{\theta} = \frac{X_1 + \cdots + X_N}{N}$$

2）中央値（メディアン）

大きさの順序で並べ替える



3）トリム平均

$$\hat{\theta} = \frac{1}{N-2m}\sum_{i=m+1}^{N-m} X_{(i)}$$

$\begin{Bmatrix} \text{小さいもの}m\text{ケ} \\ \text{大きいもの}m\text{ケ} \end{Bmatrix} \rightarrow$ 使わない

4）加重平均

$$\hat{\theta} = \underbrace{C_1X_1 + \cdots + C_NX_N}_{\text{特殊な場合，中央値，トリムを含む}} \qquad \text{但し，} C_1 + \cdots + C_N = 1$$

5）幾何平均

$$\hat{\theta} = (X_1 \cdots X_N)^{1/N}$$

推定量の不偏性

$\hat{\theta} = \hat{\theta}(X_1, \cdots, X_n)$

$E(\hat{\theta}|\theta) = \theta$ →この時「$\hat{\theta}$ は不偏推定である」という

k 通りの不偏推定量 $\hat{\theta}_1, \cdots, \hat{\theta}_k$ があったとする.

そして，それらの分散がすべて σ^2 に一致しているとする.

$V(\hat{\theta}_1) = V(\hat{\theta}_2) = \cdots = V(\hat{\theta}_k) = \sigma^2$

この時，次のような推定量 $\hat{\theta}^*$ を定義

$$\hat{\theta}^* = \frac{1}{k}\sum_{i=1}^{k}\hat{\theta}_i$$

k 通りの推定方法を組み合わせて，よりよい（より分散の小さい）推定量が得られる.

$\hat{\theta}^*$ は次の性質を有する.

$$\begin{cases} 1.\text{不偏性} & E(\hat{\theta}^*) = \theta \\ 2.\text{分散} & V(\hat{\theta}^*) \le \sigma^2 \end{cases}$$

証明）

$$E(\hat{\theta}^*) = \frac{1}{k}\sum_{i=1}^{k}E(\hat{\theta}_i) = \theta$$

$$E(\hat{\theta}_i) \overset{\text{不偏性}}{=} \theta$$

$V(\hat{\theta}^*)$ のため $\sum_{i=1}^{k}(\hat{\theta}_i - \theta)^2$ を計算する.

$$\sum_{i=1}^{k}(\hat{\theta}_i - \theta)^2 = \sum_{i=1}^{k}(\hat{\theta}_i - \hat{\theta}^* + \hat{\theta}^* - \theta)^2$$

$$= \sum_{i=1}^{k}\left\{(\hat{\theta}_i - \hat{\theta}^*)^2 + 2(\hat{\theta}_i - \hat{\theta}^*)(\hat{\theta}^* - \theta) + (\hat{\theta}^* - \theta)^2\right\}$$

$$\boxed{\begin{array}{l} \hat{\theta}^* \overset{\Delta}{=} \frac{1}{k}\sum_i \hat{\theta}_i \\ \sum_{i=1}^{k}(\hat{\theta}_i - \hat{\theta}^*) = \sum_{i=1}^{k}\hat{\theta}_1 - k\hat{\theta}^* = 0 \end{array}}$$

$$= \sum_{i=1}^{k}(\hat{\theta}_i - \hat{\theta}^*)^2 + k \times E\left[(\hat{\theta}^* - \theta)^2\right] \ge kE\left[(\hat{\theta}^* - \theta)^2\right]$$

$$\Rightarrow k\sigma^2 \ge kE\left[(\hat{\theta}^* - \theta)^2\right] \Rightarrow \sigma^2 \ge E\left[(\hat{\theta}^* - \theta)^2\right] = V(\hat{\theta}^*)$$

$\hat{\theta} = \hat{\theta}(X_1, \cdots, X_N)$: 母数 θ の不偏推定量

この時 $X_1, X_2, \cdots, X_N$ の対称関数として表される不偏推定量 $\hat{\theta}^*$ で，

$\underline{V(\hat{\theta}^*)} \le V(\hat{\theta})$ となるものが存在する.

平均値の不偏推定

$X_i = \theta + \varepsilon_i$

θ :真の未知パラメタ

$\{\varepsilon_i\}$の分布が分かっていない場合

$\hat{\theta}^* = \frac{1}{N}\sum_{i=1}^{N} X_i$ （算術平均）

θに対する一様最小分散不偏推定量

算術平均以外の不備推定量で$X_1, \cdots, X_N$について対称なものがあったとする

$T \overset{\Delta}{=} \hat{\theta}^+ - \hat{\theta}^*$と定義

$E(T) = 0$

任意の分布に対してこれを満足するTは$T = 0$しかあり得ない.