確率と統計(０)
「統計的推定(第１１章)」

- 担当教員：　杉山　将（計算工学専攻）
- 居室：　W8E-505
- 電子メール:　　　sugi@cs.titech.ac.jp
- 授業のウェブサイト：

  http://sugiyama-www.cs.titech.ac.jp/~sugi/

母集団の推定

- 一般に確率分布には**母数**(parameter)が含まれており，母数の値によって分布の形状が異なる
  - 例) 正規分布の場合，期待値 μ と分散 σ^2:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right)$$

- 標本が従う確率分布の種類がわかっていても母数の値を決めなければ分布の形状は決まらない
- **統計的推定**(statistical inference)：母分布 (の母数) を標本から推定する

推定法の種類

- **点推定**(point estimation)：真のパラメータ値をある一つの推定値で予想する方法
  - **モーメント法**(method of moments)
  - **最尤法**(maximum likelihood estimation)

  推定には何らかの誤差が伴うため，別個，誤差評価を行なう必要がある．

- **区間推定**(interval estimation)：推定には誤差が伴うことを考慮し，真のパラメータ値が入る確率がある値（例えば９９%）以上である範囲を求める方法

統計的推定の用語

- **推定量**(estimator)：母数を推定した量．標本 $\{x_i\}_{i=1}^n$ の関数であり，確率変数である．
  - 推定量はハットをつけて表すことが多い．

$$\hat{\mu} \qquad \hat{\sigma}^2$$

- **推定値**(estimate)：推定量に標本の具体的な値を代入した値

- 推定量は確率変数なので，確率分布を持つ．また，推定量の期待値や分散を考えることもできる．

最尤（さいゆう）法

- **尤度**(likelihood)：手元にある標本が得られる確率
- **最尤原理**：現実の標本は確率最大のものが実現したと考える原理
- **最尤法**(maximum likelihood estimation)：尤度を最大にするパラメータの値を推定値とする方法

最尤とは「もっとももっともらしい」という意味

最尤法 (続き)

- **例)** 成功確率が p のベルヌーイ分布に従う標本

{成功、失敗、成功、成功、成功}

が与えられたとする. 尤度は

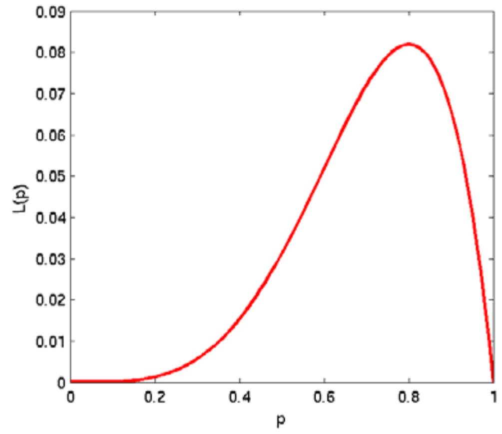
$$L(p) = p^4(1-p)$$

このとき,

$$L'(p) = p^3(4-5p)$$

なので,

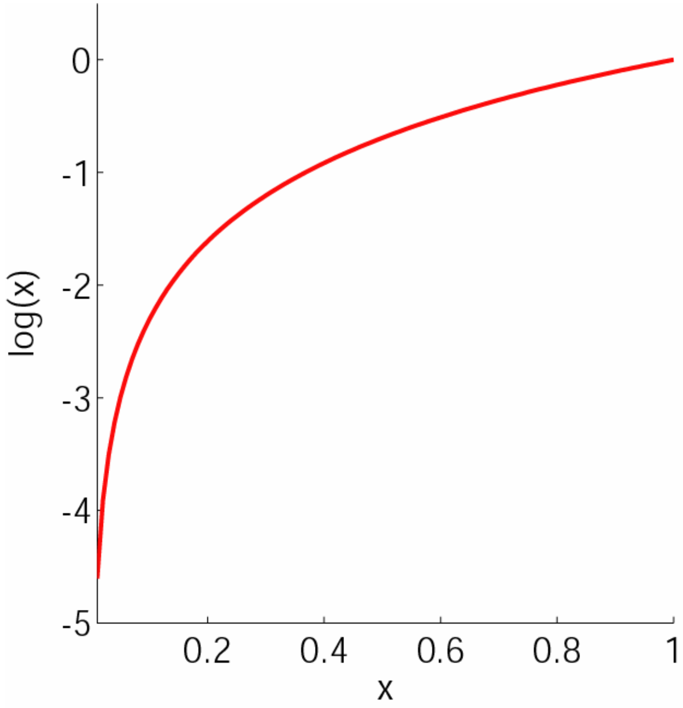
$$\widehat{p}_{ML} = \operatorname*{argmax}_p L(p) = 0.8$$



- 5回中4回が成功なので, 0.8は最も尤もらしいであろう.

最尤法 (続き)

- **対数尤度**(log-likelihood)：尤度の対数
- 対数は単調増加関数なので，対数を取っても尤度の大小関係は変わらない．対数尤度を最大にするようにパラメータを決めたほうが計算が簡単になることがある．



演習

- 正規分布 $N(\mu, \sigma^2)$ に従うi.i.d.標本 $\{x_i\}_{i=1}^n$ の尤度は

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} f(X_i)$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right)$$

1. 対数尤度を求めよ.
2. 期待値 μ の最尤推定量 $\widehat{\mu}_{ML}$ を求めよ.
3. 分散 σ^2 の最尤推定量 $\widehat{\sigma}^2_{ML}$ を求めよ.

推定量の良さの尺度

- 最尤法によって得られる推定量は，果たして本当のパラメータ値の良い推定量なのだろうか？

- 推定量の良さを評価するために次の尺度がよく用いられる．
  - **不偏性**(unbiasedness)
  - **有効性**(efficiency)

不偏性

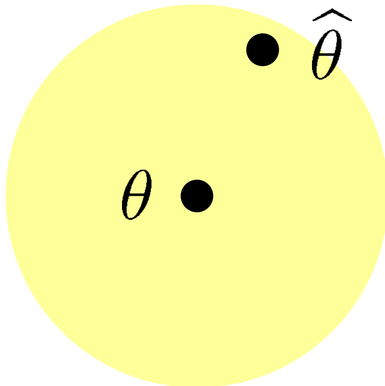
- **不偏推定量**(unbiased estimator)：**期待値が本当の値と一致する推定量**

$$E[\widehat{\theta}] = \theta$$

  - **注意：ここでの期待値は，n 個の標本全てに関する期待値である**

$$E[Z] = \int\int\cdots\int Zf(X_1)f(X_2)\cdots f(X_n)dX_1dX_2\cdots dX_n$$

- **不偏推定量は真の値の周りに偏りなく分布するため，「クセ」のない素直な推定量である．**



不偏推定量の例

- 正規分布の期待値 μ の最尤推定量 $\widehat{\mu}_{ML}$ は不偏推定量である.

$$E[\widehat{\mu}_{ML}] = \mu \qquad \widehat{\mu}_{ML} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- **証明:**

$$E[\widehat{\mu}_{ML}] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$$

不偏推定量の例 (続き)

- 正規分布の分散 σ^2 の最尤推定量 $\widehat{\sigma}^2_{ML}$ は**不偏でない**.

$$\widehat{\sigma}^2_{ML} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \widehat{\mu}_{ML})^2$$

$$E[\widehat{\sigma}^2_{ML}] \neq \sigma^2$$

$$\widehat{\mu}_{ML} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- 以下の推定量が不偏である:

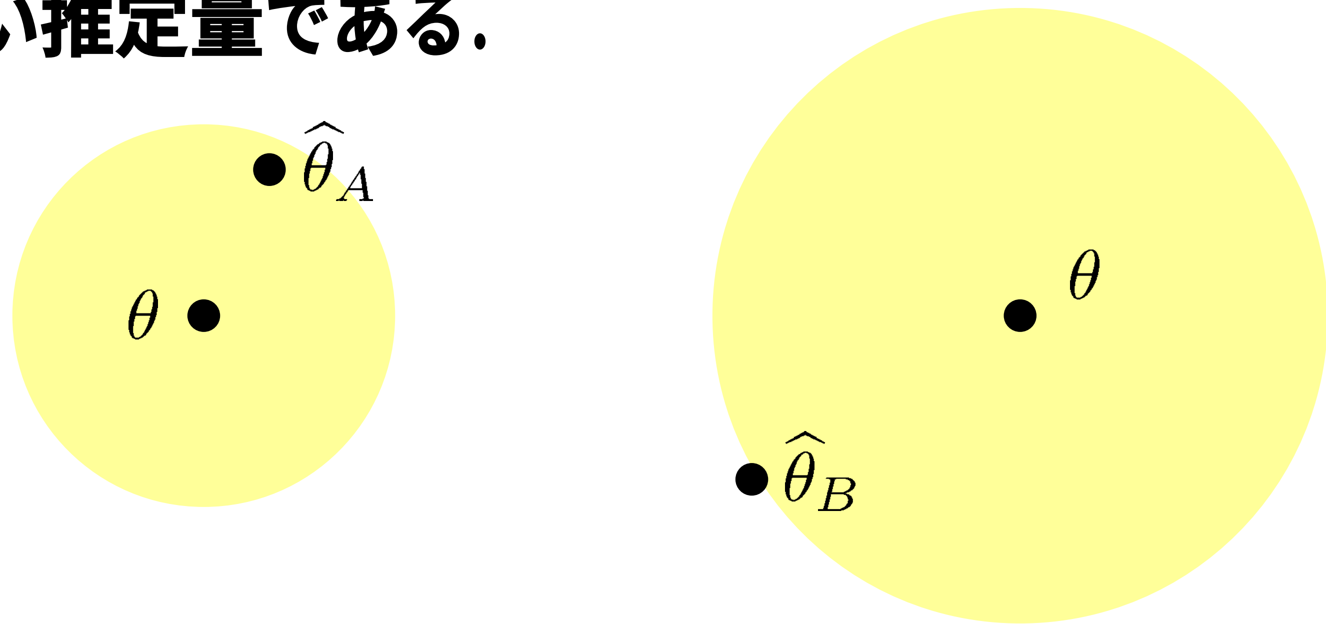
$$\widehat{\sigma}^2_{U} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \widehat{\mu}_{ML})^2$$

$$E[\widehat{\sigma}^2_{U}] = \sigma^2$$

- **証明は宿題！**

有効性

- **有効推定量**(efficient estimator)：**不偏推定量の中で分散が最小の推定量**
- **推定量が不偏でも，分散が大きければ不安定である．**
- **有効推定量は，偏りがなくさらに散らばりも小さいため，好ましい推定量である．**



- **注意**：**分散だけを評価しても無意味**
  - $\widehat{\theta}_C = 0$ **は分散ゼロ！**

漸近有効性

- 有効推定量を構成することは必ずしも容易ではない.
- 漸近有効推定量(asymptotic efficient estimator):標本数が十分大きいときに偏りがなく分散が最小
- 最尤推定量は一般に漸近有効推定量である!

- 注意:上記の漸近有効推定量の定義は厳密ではない.正確には,漸近正規推定量(分布が漸近的に正規分布になる推定量)の中でクラメル·ラオの下限を漸近的に達成する推定量を漸近有効推定量という.

宿題

1. 正規分布の分散の最尤推定量は**不偏でない**ことを証明せよ．

$$E[\widehat{\sigma}^2_{ML}] \neq \sigma^2$$

ヒント：

$$\widehat{\sigma}^2_{ML} = \frac{n-1}{n^2}\sum_{i=1}^{n} X_i^2 - \frac{1}{n^2}\sum_{i\neq j}^{n} X_i X_j$$

$$\sigma^2 = \mu_2 - \mu_1^2$$

2. 次の推定量が，正規分布の分散の不偏推定量であることを証明せよ．

$$\widehat{\sigma}^2_U = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \widehat{\mu}_{ML})^2$$

$$\widehat{\mu}_{ML} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

宿題 (続き)

3. 計算機実験:

- Octaveなどを用いて，期待値5，分散2の正規分布から標本を生成し，期待値と分散の最尤推定値を計算せよ．
- $n = 10$ に対して上記の実験を100回行ない，**期待値の最尤推定量のヒストグラム**を図示せよ．そして，**不偏性**が成り立つことを実験的に確認せよ．
- 上記の実験で n を増やしていったときの**期待値と分散の最尤推定値のグラフ**を図示せよ．そして，一致性が成り立つことを実験的に確認せよ．

試験について

- 日時：7月27日 (金) 3,4限
- 場所：S125
- 試験内容：
  - 専門用語の英語名
  - 記述問題 (選択式2問)
- 教科書, ノートは持ち込み不可！