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，平均 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．例えば
  - **モーメント法**(method of moments)
  - **最尤法**(maximum likelihood estimation)

  がある．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

点推定

- 以下, i.i.d.標本 $\{x_i\}_{i=1}^n$ からその母集団の母数を**推定**(estimation)する問題を考える
- **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, 推定量も確率変数である．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

モーメント法

- r次のモーメント $\mu_r = E[X^r]$ を標本平均で推定する：

$$\widehat{\mu}_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r$$

- **モーメント法**：モーメントを通して母集団に関する情報を求める方法
- **例)** 母集団が正規分布 $N(\mu, \sigma^2)$ に従うとき，

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

より

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

最尤法

- **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．最尤とは「もっとももっともらしい」という意味である．
- **最尤法**：尤度を最大にするパラメータの値を推定値とする方法
- **対数尤度**(log-likelihood)：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．実際には，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多い．

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)

- 例) 正規分布 $N(\mu, \sigma^2)$ に従うi.i.d.標本 $\{x_i\}_{i=1}^n$ が与えられたとする．尤度は

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

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

対数尤度は

$$\begin{aligned}\log L(\mu, \sigma^2) &= \sum_{i=1}^n \log f(X_i) \\ &= -\frac{n\log(2\pi)}{2} - n\log\sigma - \sum_{i=1}^n \frac{(X_i-\mu)^2}{2\sigma^2}\end{aligned}$$

最尤法 (続き)

- 対数尤度の偏微分は

$$\frac{\partial \log L}{\partial \mu} = \sum_{i=1}^{n} \frac{(X_i - \mu)}{\sigma^2}$$

$$\frac{\partial \log L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{2\sigma^4}$$

- 対数尤度の偏微分をゼロとおくと

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

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

$\Longrightarrow$ モーメント法と同じ！

推定量の良さの尺度

- モーメント法や最尤法によって何らかの推定量が得られる．それら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)
  - **一致性**(consistency)
  - **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

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

不偏推定量の例

- 例) 正規分布の平均 μ の最尤推定量 $\widehat{\mu}_{ML}$ は不偏推定量である.

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

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

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

$$E[Z] = \int\int\cdots\int Z f(X_1) f(X_2)\cdots f(X_n) dX_1 dX_2 \cdots dX_n$$

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.

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

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

- 正規分布の分散は, 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！

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

一致性

- **一致推定量**(consistent estimator)：標本数 n を大きくしていったとき，本当の値に収束する推定量
- 正確には，全ての $\varepsilon > 0$ に対して，$n \to \infty$ のときに
$$P(|\widehat{\theta}_n - \theta|) > \varepsilon) \to 0$$
となる推定量 $\widehat{\theta}_n$ を θ の一致推定量であるという．
- 十分にたくさん標本があるとき，良い推定量であることが保証される．

一致推定量の例

- 正規分布の μ, σ^2 の最尤推定量 $\widehat{\mu}_{ML}, \widehat{\sigma}^2_{ML}$ は一致推定量である.
- 証明は, 大数の法則から明らか.

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．**
- **しかし，有効推定量を構成することは必ずしも簡単ではない．そこで条件を少し緩めた漸近有効性**(asymptotic efficiency)**がよく用いられる．**
- **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！**

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. **以下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.**

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

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