応用統計学 2017 第8回 漸近正規性，信頼区間

2017 年 11 月 29 日 (水)
清 智也 sei@mist.i.u-tokyo.ac.jp
http://ur0.pw/yTzt

最尤推定量の利点として，一致性，漸近正規性，漸近有効性が挙げられる。特に漸近正規性によって推定量の近似的な信頼区間を求めることができる。

- 一致性，漸近正規性，漸近有効性，標準誤差，信頼区間，区間推定[1]。
- 「標準正規分布の上側 2.5%点」$= 1.96$。正確には $1.959964\cdots$

演習問題

問題 8-1 (区間推定の基礎)**.** $X_1, \ldots, X_n$ は独立で，平均 μ，分散 σ^2 の正規分布に従うものとする。また σ^2 の値は既知とする。このとき，区間

$$\left[\bar{X} - \frac{1.96\sigma}{\sqrt{n}}, \bar{X} + \frac{1.96\sigma}{\sqrt{n}}\right]$$

は μ の 95%信頼区間であることを示せ。ただし $\bar{X}$ は $X_1, \ldots, X_n$ の標本平均を表す。

問題 8-2 (中心極限定理の復習)**.** $X_1, \ldots, X_n$ は独立で，$[0, 1]$ 上の一様分布に従うものとする。また

$$Z_n = \left(\sum_{i=1}^{n} \cos(2\pi X_i)\right)^2 + \left(\sum_{i=1}^{n} \sin(2\pi X_i)\right)^2$$

とおく。$n \to \infty$ とすると，Z_n/n の分布は平均 1 の指数分布に収束することを示せ。

問題 8-3. $X_1, \ldots, X_n$ は平均 μ，分散 σ^2 の独立同一分布に従う確率変数とする。

(i) 標本平均を $\bar{X}$ とおく。$\sqrt{n}(\bar{X} - \mu)$ の漸近分布を求めよ。

(ii) μ の 95%信頼区間を求めよ[2]。ただし σ^2 は未知とする。

注意：指数型分布族の場合を除けば，$\bar{X}$ は一般には μ の最尤推定量にはならず，漸近有効とは限らない。

問題 8-4. $X_1, \ldots, X_n$ は確率 p のベルヌーイ試行に従うものとする。

(i) 最尤推定量を $\hat{p}$ とおく。$\sqrt{n}(\hat{p} - p)$ の漸近分布を求めよ。

(ii) $\hat{p}$ の 95%信頼区間を求めよ。

---

[1] consistency, asymptotic normality, asymptotic efficiency, standard error, confidence interval, interval estimation.

[2] 正確には，「漸近分布に基づく近似的な 95%信頼区間」という意味である。以下同様である。なお，母集団が正規分布に従う場合には，t 分布に基づく正確な 95%信頼区間を求めることができる。詳しくは次回説明する。

問題 8-5 (難しい). 指数型分布族について，最尤推定量の一致性と漸近正規性の証明の流れを確認する[3]。$X_1, \ldots, X_n$ は独立で，それぞれ確率密度関数 $f(x;\theta) = a(x)e^{\theta s(x) - \psi(\theta)}$ $(\theta \in \mathbb{R})$ の分布に従うとする。また，$\mu(\theta) = \psi'(\theta)$ とおき，θ のフィッシャー情報量を $I(\theta)$ とおく。

(i) $I(\theta) = \psi''(\theta) = \mu'(\theta)$ を示せ（問題 7-5 の復習）。

(ii) 最尤推定量 $\hat{\theta}$ は $\mu(\hat{\theta}) = \bar{s}$ の解であることを示せ。ただし $\bar{s} = n^{-1}\sum_{t=1}^{n} s(X_t)$ とする。

(iii) $s(X_1)$ の期待値と分散はそれぞれ $\mu(\theta)$，$I(\theta)$ であることを示せ。

(iv) 大数の法則を用いて $\bar{s}$ が $\mu(\theta)$ に確率収束することを示せ。

(v) $\hat{\theta}$ は一致性を持つこと，つまり $\hat{\theta}$ は θ に確率収束することを示せ。ただし次の補題を用いてよい。

> 補題　Y_n は確率変数列とし，ある定数 $c \in \mathbb{R}$ に確率収束すると仮定する。このとき，任意の連続関数 $h(y)$ に対して，$h(Y_n)$ は $h(c)$ に確率収束する。

(vi) 中心極限定理を用いて $\sqrt{n}(\bar{s} - \mu(\theta))$ が $\mathrm{N}(0, I(\theta))$ に分布収束することを示せ。

(vii) $\sqrt{n}(\hat{\theta} - \theta)$ は $\mathrm{N}(0, I(\theta)^{-1})$ に分布収束することを示せ。ただし次の定理を用いてよい。

> 定理（デルタ法）　Y_n は確率変数列とし，ある定数 $c \in \mathbb{R}$ と $\sigma^2 > 0$ が存在して，$\sqrt{n}(Y_n - c)$ が $\mathrm{N}(0, \sigma^2)$ に分布収束すると仮定する。このとき，任意の C^1 級関数 $h(y)$ に対して，$\sqrt{n}(h(Y_n) - h(c))$ は $\mathrm{N}(0, h'(c)^2\sigma^2)$ に分布収束する。

宿題 8

問題 8-6. アンケートのデータに基づき，学生の平均睡眠時間の 95%信頼区間を求めよ。

[3] 指数型分布族に限らない一般の場合の証明については，例えば竹村彰通「現代数理統計学」創文社の 13 章を参照せよ。