

数理統計学まとめ (その8): 第5章 推定

1 点推定

母集団の特性を表す値 (母集団の平均や分散など) を母数 (またはパラメータ) と呼ぶ。これらは標本の平均や分散と区別するため母平均や母分散と呼び, その値は μ, σ^2 などのギリシャ文字で表す。母相関係数も ρ と書き, 区別する。母数の推定のために用いる統計量を推定量という。推定量は $\hat{\cdot}$ をつけて $\hat{\mu}$ や $\hat{\sigma}^2$ と表す。

モーメント法による点推定

母数の推定する方法としてまずモーメント法がある。モーメント法は理論的には単純で, たとえば母平均 μ を推定するには標本平均 \bar{X} を使い, 母分散を推定するには標本分散 S^2 を用いる事になる。 $\hat{\mu} = \bar{X}$ で母平均の推定量として標本平均を用いた事を表している。分散の推定量に標本分散をとったときは $\hat{\sigma}^2 = S^2$ と書く。

2 推定量とその性質

推定量がどのくらいよく母数を推定しているかを測る基準として不偏性, 有効性, 一致性がある。推定する母数は θ と書き, 推定量を $T = T(X_1, \dots, X_n)$ と書くことにする。

不偏性 この推定量の平均 (T は確率変数。平均は母集団の分布でとる) は母数 θ と一致する。

$$E(T) = E(T(X_1, \dots, X_n)) = \theta$$

有効性 T の分散はできるだけ小さいことが望ましい (標本をとったときばらつきが小さい)。未知母数 θ の二つの不偏推定量 T_1, T_2 を比べたとき, $V(T_1) < V(T_2)$ ならば T_1 は T_2 より有効であるという。

$T = T(X_1, \dots, X_n)$ が θ の不偏推定量のとき, X_i たちの共通の分布密度関数を $f_\theta(x)$ と書くとき,

$$V(T) \geq \frac{1}{nE\left(\left(\frac{\partial \log f_\theta}{\partial \theta}\right)^2\right)} =: \sigma_0^2 \quad (\text{Cramér-Rao の不等式})$$

が成り立つ． $V(T) = \sigma_0^2$ を満たす不偏推定量を有効推定量という．

$$I(\theta) := E \left(\left(\frac{\partial \log f_\theta(X)}{\partial \theta} \right)^2 \right)$$

は Fisher の情報量 と呼ばれる．

一致性 母数 θ の推定量 T は標本の大きさによるので，その依存性を明示するとき T_n と書く．任意の $\varepsilon > 0$ に対して

$$P(|T_n - \theta| > \varepsilon) \rightarrow 0 (n \rightarrow \infty)$$

が成り立つとき T は 一致推定量 という．

例 2.1 母平均 μ の推定量としての標本平均 \bar{X} について考えてみる．これまでに行ったことから $E(\bar{X}) = \mu$ なので， \bar{X} は不偏性をもつ．また，大数の法則が成り立つとき（母分散が有限なら成り立つ）

$$P(|\bar{X} - \mu| > \varepsilon) \rightarrow 0 \quad (n \rightarrow \infty)$$

となるので，一致性も成り立つ．有効性を見るには母集団の分布密度関数が必要なので， μ が未知母数という意味で θ を代わりに使い，正規分布 $N(\theta, \sigma^2)$ のときを考えると，

$$f_\theta(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\theta)^2}{2\sigma^2}}$$

となる．これから

$$\frac{\partial \log f_\theta(x)}{\partial \theta} = -\frac{\theta - x}{\sigma^2}$$

となり，

$$E \left(\left(\frac{\partial \log f_\theta(x)}{\partial \theta} \right)^2 \right) = E \left(\frac{(\theta - X)^2}{\sigma^4} \right) = \frac{1}{\sigma^2}$$

を得る．これより， $\sigma_0^2 = \frac{\sigma^2}{n}$ と計算できるが， $V(\bar{X}) = \frac{\sigma^2}{n}$ だったので， \bar{X} は有効な推定量にもなっている．

例 2.2 X_1, \dots, X_n は母平均 μ ，母分散 σ^2 の母集団からの無作為標本とする．標本分散

$$S^2 = \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2 = \frac{1}{n} \sum_{j=1}^n (X_j - \mu)^2 - (\bar{X} - \mu)^2$$

について見てみよう．母平均が μ ，母分散が σ^2 なので

$$\begin{aligned} E(S^2) &= \frac{1}{n} \sum_{j=1}^n E((X_j - \mu)^2) - E((\bar{X} - \mu)^2) \\ &= \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2 \end{aligned}$$

となり， S^2 は不偏推定量ではない．不偏標本分散

$$U^2 = \frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X})^2$$

がこれを修正して母分散の不偏推定量となっている．母集団の分布が $N(\mu, \sigma^2)$ のとき， $(nS_n^2)/\sigma^2$ は自由度 $n-1$ のカイ 2 乗分布に従い，その分散は $2(n-1)$ だったので，

$$V(S_n^2) = \frac{\sigma^4}{n^2} V\left(\frac{nS_n^2}{\sigma^2}\right) = \frac{\sigma^4}{n^2} 2(n-1), \quad V(U_n^2) = \frac{2\sigma^4}{n-1}$$

となり，チェビシェフの不等式から

$$\begin{aligned} P(|S_n^2 - \sigma^2| > \varepsilon) &\leq \frac{E\left(\left(S_n^2 - \frac{n-1}{n}\sigma^2 - \frac{1}{n^2}\sigma^2\right)^2\right)}{\varepsilon^2} \\ &= \frac{1}{\varepsilon^2} \left(V(S_n^2) + \frac{\sigma^4}{n^2} \right), \\ P(|U_n^2 - \sigma^2| > \varepsilon) &\leq \frac{1}{\varepsilon^2} V(U_n^2) \end{aligned}$$

となる．どちらも $n \rightarrow \infty$ のとき 0 に収束する．したがって S_n^2 も U_n^2 も一致推定量である．

3 最尤法

推定したい母数を θ と書き，母集団の分布密度関数を $f_\theta(x)$ と書いておく．無作為標本 X_1, \dots, X_n はそれぞれ $f_\theta(x)$ を分布密度関数に持つ独立な確率変数になっており，得られた値 X_1, \dots, X_n を実現する確率密度は

$$L(\theta) = \prod_{j=1}^n f_\theta(X_j) = f_\theta(X_1) \times \cdots \times f_\theta(X_n)$$

となる． L を尤度関数と呼ぶ．この確率密度が最大になるときこの標本 X_1, \dots, X_n がもっとも得られやすいと考え， $L(\theta)$ が最大になるような θ の値 $\hat{\theta}$ を θ の最尤推定量と呼ぶ．実際には対数尤度関数 $l(\theta) = \log L(\theta)$ の最大値問題に帰着する事が多い．

例 3.1 母集団が正規分布しており，母平均は未知だが母分散が σ^2 とわかっている場合の平均の最尤推定量を求めよう．大きさ n の標本 X_1, \dots, X_n に対して対数尤度関数は

$$l(\theta) = -\frac{n}{2} \log(2\pi\sigma^2) - \sum_{j=1}^n \frac{(X_j - \theta)^2}{2\sigma^2}$$

となり，

$$l'(\theta) = -\sum_{j=1}^n \frac{\theta - X_j}{\sigma^2}$$

だから $\theta = (X_1 + \dots + X_n)/n = \bar{X}$ で $l(\theta)$ は最大になる．よってこのときの母平均の最尤推定量は $\hat{\theta} = \bar{X}$ である．