最尤推定法 (訂正版)

1　最尤推定

最尤法 (most likelyhood method, あるいは maximum likelyhood method) とは、未知パラメータを求める方法。標本データ値を取る可能性がもっとも高くなるように定める。パラメータ θ をもつ分布 $f(x|\theta)$ に対して尤度関数 $\mathrm{lik}(\theta) = f_n(x|\theta) = \Pi_{i=1}^n f(x_i|\theta), x = (x_1, x_2, \cdots, x_n)$ とおき、$\hat{\theta}(x)$ とは、最大値 $\max_\theta \mathrm{lik}(\theta)$ を与えるもので定義される。最尤推定量、MLE(maximum likelihood estimator) とは, どのような計算で求める関数かを意味するもので、関数 $\hat{\theta}(x)$ をいう。尤度の最大を求めるためには微分をして計算する。尤度方程式: $\dfrac{d\mathrm{lik}(\theta)}{d\theta} = 0$ を解いて計算すればよい。離散型分布の場合では、差分 $(\Delta\mathrm{lik}(\theta) = \mathrm{lik}(\theta) - \mathrm{lik}(\theta \pm 1))$ から最大を求める。

問1 （本文資料 10 ページ）$X \sim \mathrm{Bin}(1, p)$ ベルヌーイ分布 (1 枚のコイン投げ) ; $f_X(x) = p^x(1-p)^{1-x}, x = 0, 1$ において, n 個の標本データ $x_i, i = 1, 2, \cdots, n$ に対する p の MLE を求めよ。

問2（本文資料 12 ページ）$X \sim U(0, \theta)$ 一様分布（同じ確からしさ、乱数）; $f_X(x) = 1/\theta, 0 < x < \theta$ のとき、標本データ $x_1, x_2, \cdots, x_n$ に対する θ の MLE は $\max_i x_i$ で与えられることを示せ。

2　十分統計量

統計量 $T = T(X)$ がパラメータ θ の十分統計量 (sufficient) であるとは $T(X) = t$ として与えた X の条件付き分布が θ に依存しないとき。必要十分条件として次の定理が知られている。密度関数が分解される:

$$f(x|\theta) = g(T(x), \theta)h(x)$$

の形に表現できること。

問3（本文資料 11 ページ）2 つの確率変数 X_1, X_2 は独立で、同じ 2 項分布 $\mathrm{Bin}(n, p)$ に従うとき、(1) 条件付き確率 $P(X_1 = x|X_1 + X_2 = r)$ を求めよ。(2) $X_1 + X_2$ は p の十分統計量であることを示せ。

3　不偏推定量

未知パラメータ θ の推定量 $T(x)$ が不偏である (unbiased) であるとは、その期待値 $E[T(X)] = \theta$ が成り立つとき。推定量の期待値がちょうど推定するパラメータと等しい、ずれ（偏り）が起きていないという意味。

つぎの仮定の下で問 4、問 5 を答えよ。独立、同一分布にしたがう標本データ $X_1, X_2, \cdots, X_n$ から、これらの分布の平均 μ、分散 σ^2 の不偏推定量を求める。仮定: $EX_i = \mu$, $var(X_i) = E(X_i - \mu)^2 = \mu_2 - \mu^2 = \sigma^2, i = 1, 2, \cdots, n$。ここで 下付きインデックス $\mu_2 = E(X_i^2)$ は 2 次モーメント。

問4 （本文資料 6 ページ）定数 $a_i, i = 1, 2, \cdots, n$ より、$T(x) = \sum_i a_i x_i$ とおく。(1) $\sum_i a_i = 1$ ならば、$E[T(X)] = \mu$ すなわち、不偏推定量であることを示せ。(2) さらに分散 $\mathrm{var}(T(X))(\bigstar)$ を最小とするような定数 a_i を求めよ。★確率変数 X を $T(X)$ に訂正。

問5 （本文資料 6 ページ）2 つの推定量 $U_n = \dfrac{2}{n(n-1)} \sum_{i<j} \dfrac{(x_i - x_j)^2}{2}$ と $s_x^2 = \dfrac{1}{n-1} \sum_i (x_i - \overline{X})^2$ とおくとき、(1) $U_n = s_x^2$ が成り立ち, (2) ともに分散 σ^2 の不偏推定量である。ここで $\sum_{i<j}$ とは、整数 $1, 2, \cdots, n$ の中で, 条件 $\{i < j\}$（等号は含まれない）を満たす組 (i, j) に対する 2 重和。また $\overline{X} = \dfrac{1}{n} \sum_i x_i$ の算術平均。