STAT 391

1/21/25

Lecture 5

Tutorial on plotting data

ML estimation

Examples

Poisson

Tied parameters

Lecture Notes II – Maximum Likelihood Estimation for Discrete Distributions

Marina Meilă
mmp@stat.washington.edu

Department of Statistics
University of Washington

April, 2023

Max Likelihood Principle ✓

ML estimation for arbitrary discrete distributions ✓

Other ML estimation examples ←

ML estimate as a random variable

Reading: Ch. 4.1, 4.2

 Examples

$m = 2$

UW Student speaks Chinese $= X \in \{0,1\}$

estimate

My "guess" Model

$\theta_1^{ML} = \frac{3}{10} = 0.3 = \Pr[$a random UW student speaks Chinese$]$

$n = 10$

$n_1 = 3$

$n_0 = 7$

$\theta_0^{ML} = 0.7$

Ex 2 $m = 4$

true $m^* \gg 4$

Extra credit !

S = Species of trees on UW campus

= { Hisakura Cherry, Douglas Fir, Maple, W. Hemlock }

x = species of tree

ML estimation for arbitrary discrete distributions

$n = 20$

$n_C = 10$

$n_H = 1$

$n_D = 2$

$n_M = 7$

$\Longrightarrow$

C, M, H, D

$\theta_C^{ML} = 0.5$

$\theta_H^{ML} = 0.05$

$\theta_D^{ML} = 0.1$

$\theta_M^{ML} = 0.35$

Estimates

S = Species of trees on UW campus
= { Hisakura Cherry, Douglas Fir, Maple, W. Hemlock }

x = species of tree

$\{ P = (\theta_C, \theta_M, \theta_H, \theta_D) \}$

Model class

Model

3. Starting car gas ignition OTHER

$S = \{0, 1, 2, 3, 4, 5\}$

0: "car starts"

1: "battery"

4: alternator

nominal

error "codes"

$n = 20$

$n_0 = 10$

$n_1 = 6$

$n_2 = 1 = n_3 = \cdots$

$\theta_0^{ML} = 0.5$

$\theta_1^{ML} = 0.3 \ldots$

ML estimation for arbitrary discrete distributions

$S = \{0, 1, \ldots m-1\}$ but some $n_j = 0$

$S = \{0, 1, 2 \ldots \}$

$|S| = \infty$, n finite $\Rightarrow$ most $n_j = 0$

Idea 1 smoothing

$n_j = 0 \Longrightarrow \theta_j^{ML} = 0 \rightarrow$ choose $\tilde{\theta}_j > 0$ NOT ML

UNACCEPTABLE!

Idea 2 Get more data if possible

Idea 3 Choose model class with ~~finite~~ few parameters

Poisson, Geometric

Tie parameters

Other ML estimation examples

Poisson

$S = \{0, 1, 2, \ldots\}$

$x \in S$

$P(x) = \frac{\lambda^x}{x!} e^{-\lambda}$ $\quad \lambda > 0$

$\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{\lambda} = Z$ normalization constant

Queuing systems
$x = \#\text{queries} / \text{time unit}$

$n = 10$

$x^{1:10} = 3, 5, 3, 4, 0, 1, 1, 12, 2 = \mathcal{D}$ $\quad s = 35$

ML Estimation for λ

1. Likelihood $\quad P(\mathcal{D} \mid \lambda) \equiv L(\lambda) = \prod_{i=1}^{n} P(x^i \mid \lambda) =$

$$= \prod_{i=1}^{n} \left(\frac{\lambda^{x^i}}{x^i!} \cdot e^{-\lambda} \right) = \frac{\lambda^{\sum_{i=1}^{n} x^i}}{\prod_{i=1}^{n} (x^i!)} \cdot e^{-n\lambda}$$

1'. log-l. $\quad \ln L(\lambda) = \ln\lambda \cdot \underbrace{\sum_{i=1}^{n} x^i}_{\text{data}} - n\lambda - \sum_{i=1}^{n} \ln(x^i!)$

STAT

CALC

Other ML estimation examples Poisson (cont)

$$\max_{\lambda}:$$

$$\ln L(\lambda) = \ln\lambda \cdot \underbrace{\sum_{i=1}^{n} x^i}_{\text{data}} - n\lambda + \underbrace{\text{constant}}_{\text{indep of } \lambda} = \ell(\lambda) \qquad \lambda \in (0,\infty)$$

$$\ell'(\lambda) = 0 = \frac{1}{\lambda} \cdot \underbrace{\sum_{i=1}^{n} x^i}_{S} - n + 0 \Rightarrow \frac{S}{\lambda} = n \Rightarrow \lambda = \frac{S}{n}$$

$$\boxed{\lambda^{ML} = \frac{1}{n}\sum_{i=1}^{n} x^i}$$

ML estimator

$=$ average x

$S = 35$

$n = 10$

$\Rightarrow \lambda^{ML} = 3.5$

$S = \sum_{i=1}^{n} x^i$

SUFFICIENT STATISTIC

Other ML estimation examples Model with tied parameters

$S = \{0, 1, \cdots 6\}$ car starting ?

$\underline{n = 20}$

$n_0 = 10$

$n_1 = 6$

$n_2 = \underline{0}$

$n_3 = \underline{2}$

$n_4 = n_5 = \underline{1}$

$n_6 = 0$

Model class $\{ P = (\theta_0, \theta_1, \theta_*) \}$ 3 params.

same Probability for $x = 2, 3, 4, 5, 6$

$\theta_2 = \theta_3 = \cdots = \theta_6 = \underline{\theta_*}$

Tied params.

1. Likelihood

$$L(\theta_{0,1,*}) = \theta_0^{10} \, \theta_1^{6} \, \theta_*^{\overbrace{n_2 + n_3 + n_4 + n_5 + n_6}^{\text{suff stat}} = 4}$$

$n_0 = 10$

$n_1 = 6$

$n_* = 4$

1'. Log-L

$$\ell(\theta_{0,1,*}) = n_0 \ln \theta_0 + n_1 \ln \theta_1 + n_* \ln \theta_*$$

STAT

calc

Other ML estimation examples

$$\max_{\theta_{0,1,*}} \ \ell(\theta_{0,1,*}) = n_0 \ln\theta_0 + n_1 \ln\theta_1 + n_* \ln\theta_*$$

$\theta_{0,1,*} \in [0,1]$

s.t. $\theta_0 + \theta_1 + \theta_* \cdot 5 = 1$ $(*)$

Lagrange multiplier

$$f = n_0 \ln\theta_0 + n_1 \ln\theta_1 + n_* \ln\theta_* - \lambda(\theta_0 + \theta_1 + 5\theta_* - 1)\ (*)$$

$$\frac{\partial f}{\partial\theta_0} = \frac{n_0}{\theta_0} - \lambda 1 = 0 \Rightarrow \theta_0 = \frac{n_0}{\lambda} \qquad \theta_1 = \frac{n_1}{\lambda}$$

$$\frac{\partial f}{\partial\theta_1} = \frac{n_1}{\theta_1} - \lambda = 0 \qquad \theta^* = \frac{n_*}{5\lambda}$$

$$\frac{\partial f}{\partial\theta_*} = \frac{n_*}{\theta_*} - \lambda 5 = 0$$

$$\frac{\partial f}{\partial\lambda} = (*) = \frac{n_0}{\lambda} + \frac{n_1}{\lambda} + 5\frac{n_*}{5\lambda} = 1 \Rightarrow$$

$$\lambda = n_0 + n_1 + n_* = n$$

$$\boxed{\theta_0^{ML} = \frac{n_0}{n} \qquad \theta_1^{ML} = \frac{n_1}{n} \qquad \theta_*^{ML} = \frac{n_*/5}{n}}$$

Other ML estimation examples

How much more data?

$S = \{0, 1\}$ $\quad \theta_0 >> \theta_1$

w.p. $1-\delta$ $\quad \delta = 10^{-3}$

Q $\quad n = ?$ So that $n_1 \geq 1$ w.h.p

with high probability

Q' $\quad n$ known

θ_1 known

$P[n_1 = 0] = ?$ solve for n

$P[n_0 = n] = \theta_0^n = \boxed{(1-\theta_1)^n = \delta}$

$n \ln(1-\theta_1) = \ln \delta \Rightarrow \boxed{n^* = \frac{\ln \delta}{\ln(1-\theta_1)} = \frac{-\ln \frac{1}{\delta}}{\theta_1}}$

$\ln(1+z) \approx z$

$\ln(1-\theta_1) \approx -\theta_1$

0. $n > n^* \Rightarrow P[n_0 = n] \leq \delta$

1. $n^* \propto \frac{1}{\theta_1}$ intuitive

2. $n^* \propto \ln \frac{1}{\delta}$ $\nearrow$ when larger $1-\delta$ confidence wanted

~~Other ML estimation examples~~ What next?

Idea 1 $n_j = 0$ what to do? ← ≈next week

ML estimator as r.variable ← next lecture