STAT 391

1/23/25

Lecture 6

ML with censored data

ML estimate as r.v.

HW 1 due

HW 2 out.

Campus-trees out.

Plotting tutorial ✓

Lecture Notes II – Maximum Likelihood Estimation for Discrete Distributions

Marina Meilă
mmp@stat.washington.edu

Department of Statistics
University of Washington

January 2025

Max Likelihood Principle ✓

ML estimation for arbitrary discrete distributions ✓

· finite S · tied parameters
· Poisson

Other ML estimation examples ← · censored data

ML estimate as a random variable ←

Reading: Ch. 4.1, 4.2

Other ML estimation examples

Censored data

Ex: 1) clinical trial

Observed

$X_t = 0$

$X_{t+1\,\text{year}} = 1$

time

X t^*

not observed

2) X = color image (unobserved)

Y = B/W image observed

2') Y = low resolution (X)

X high — " — unobserved

$X \sim P_X$? $\xrightarrow{\text{deterministic}}$ $Y(X)$

unobserved observed

3) $X \in \{0, 1, 2, \dots\} = S$ $\quad P_X = \text{expo}\,(\lambda)$?

sample $X^1, X^2, \dots X^n \sim P_X$ iid unobserved

$y^i = x^i \bmod 2 \iff y^i = \begin{cases} 1 & x^i \text{ odd} \\ 0 & x^i \text{ even} \end{cases}$

Other ML estimation examples - censored data - cont

3) $x \in \{0, 1, 2, \ldots\} = S$ $\quad P_X = \mathrm{geom}(\lambda)$

sample $x^1, x^2, \ldots x^n \sim P_X$ iid unobserved

$y^i = x^i \bmod 2 \iff y^i = \begin{cases} 1 & x^i \text{ odd} \\ 0 & x^i \text{ even} \end{cases}$

ML Principle

$$L(\lambda \mid y^{1:n}) = P(y^{1:n} \mid \lambda) = \prod_{i=1}^{n} P(y^i \mid \lambda)$$

observed

Find $P[\text{observed} \mid \text{parameters}] \equiv L$

$x \in S$

$P(x) = (1-\lambda)\lambda^x$

$\lambda \in (0,1)$

$$P[Y=0 \mid \lambda] = P[x \text{ even} \mid \lambda] =$$

$$= P[0] + P[2] + \cdots P[2j] + \cdots$$

$$= \sum_{j=0}^{\infty} (1-\lambda)\lambda^{2j} = (1-\lambda)\frac{1}{1-\lambda^2} = \boxed{\frac{1}{1+\lambda}}$$

Other ML estimation examples - censored data (cont.)

$$P[y=1\mid\lambda] = \sum_{j=0}^{\infty} P[2j+1] = \cdots$$

$$= 1 - P[y=0\mid\lambda] = 1 - \frac{1}{1+\lambda} = \frac{\lambda}{1+\lambda} < P[y=0\mid\lambda]$$

$$\underline{L}(\lambda) = \prod_{j=1}^{n} \frac{\lambda^{y^i}}{1+\lambda} = \frac{\lambda^{\sum y^i}}{(1+\lambda)^n}$$

$$\theta_0 \leftarrow \left[\underset{y^i=0}{\frac{1}{1+\lambda}}\right] \qquad \left[\underset{y^i=1}{\frac{\lambda}{1+\lambda}}\right] \rightarrow \theta_1$$

log-L $\quad \ell(\lambda) = \sum_{i=1}^{n} y^i \cdot \ln\lambda - n\ln(1+\lambda)$

$$n_1 = \#\{y^i = 1\}$$

STAT

CALC

max $\ell(\lambda)$:

$\lambda \in (0,1)$

$$\ell'(\lambda) = \frac{n_1}{\lambda} - \frac{n}{1+\lambda} = 0 \qquad \frac{1+\lambda}{\lambda} = \frac{n}{n_1} \quad n = n_1 + n_0$$

$$\frac{1}{\lambda} + 1 = \frac{n_0}{n_1} + 1 \Rightarrow \boxed{\lambda^{ML} = \frac{n_1}{n_0}}$$

Other ML estimation examples - censored data (cont)

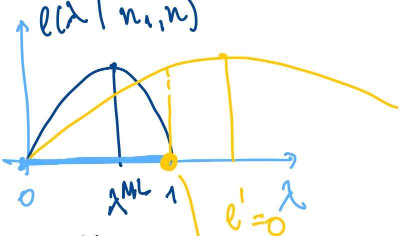
always ✓

>0

$\lambda^{ML} = \frac{n_1}{n_0}$ <1 can be >1 !! $\rightarrow$ if $\frac{n_1}{n_0} > 1 \Rightarrow \underset{\lambda \in [0,1]}{\text{argmax}}\, \ell(\lambda) = 1 = \underbrace{\lambda^{ML}}$

can't define a geom (λ)

$\lambda < 1$!!



$n_1 < n_0$

$n_1 \geq n_0$ $\quad \lambda^{ML} = 1$

get more data !

Ex given λ

$P[n_1 > n_0 \mid \lambda, n] = ?$

ML estimate as a random variable

<u>Ex</u> Tossing a coin $\theta_1 = Pr[1] = ?$ $S = \{0, 1\}$

0.5 maybe

$n = 10 \Rightarrow \underline{\underline{\theta_1^{ML}}} = \frac{n_1}{n} \neq \theta_1^{true}$

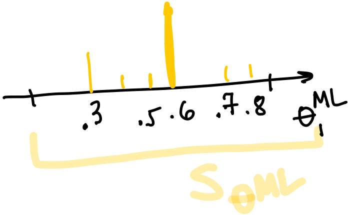
A GUESS !! $\Rightarrow$ 1. random $n = 10$

2. $S_{\theta^{ML}}^{n=10} = \{0, 0.1, 0.2, \ldots 0.9, 1\}$
$= \{ \frac{k}{n}, k = 0:n \}$

$|S_{\theta^{ML}}| = n + 1$

$S_{\theta^{ML}}^{n=100} = \{0, 0.01, \ldots .99, 1\}$

$\theta^{true} \notin S_{\theta^{ML}}$?

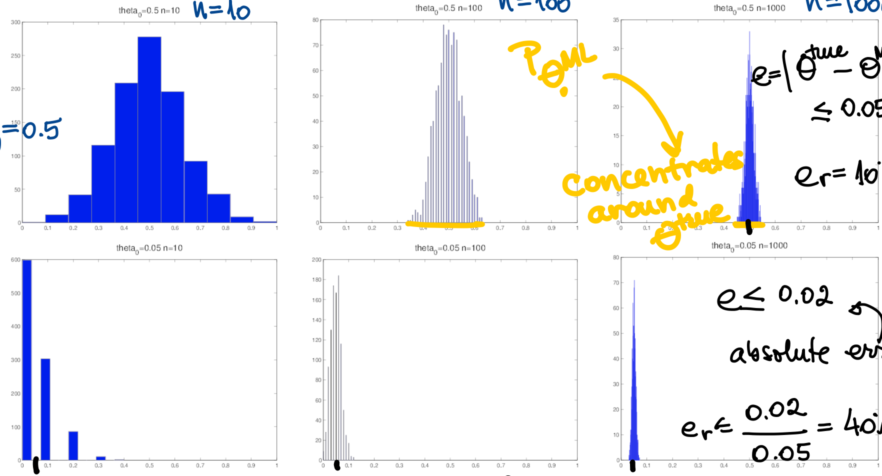


3. $P[\theta_1^{ML} = \frac{k}{n}] = ?$

Randomness of θ_1^{ML}

ML estimate as a random variable

Randomness experimentally 1000 repetitions



$n=10$ $n=100$ $n=1000$

$\theta_0 = 0.5$

$P_{\theta^{ML}}$ concentrates around θ^{true}

$e = |\theta^{true} - \theta^{ML}| \leq 0.05$

$e_r = 10\%$

$e \leq 0.02$ absolute err

$e_r \leq \dfrac{0.02}{0.05} = 40\%$

$\theta_0 = 0.05$

Relative err $= \dfrac{e}{\theta^{true}}$

ML estimate as a random variable

Randomness Theoretically

$\theta_1^{true} = .7$

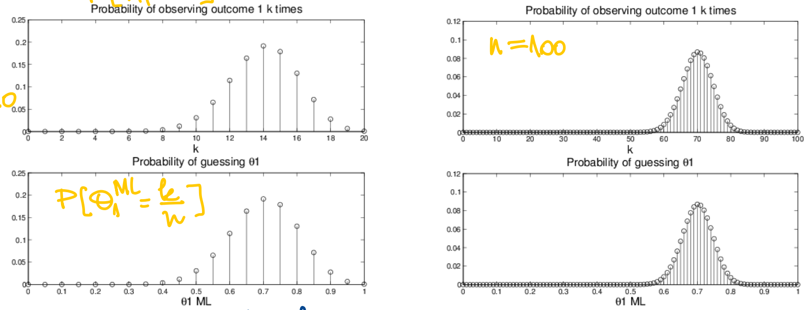
$$\binom{n}{n_0\ n_1\ \dots n_{m-1}} = \frac{n!}{n_0!n_1!\dots n_{m-1}!}$$

$P[n_1 = k]$

$n=20$

$P[\theta_1^{ML} = \frac{k}{n}]$

$n=100$



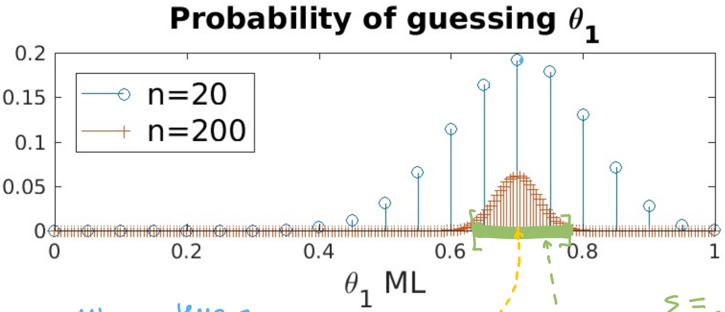
Exact $P[\theta_1^{ML} = \frac{k}{n}]$ for some $k = 0{:}n$

$S_{\theta_1^{ML}}$

$$P\left[\frac{k}{n}\right] = P[n_1 = k \mid \theta_1] = \binom{n}{n_1} \theta_1^{k} (1-\theta_1)^{n-k}$$

multinomial

ML estimate as a random variable



$\varepsilon = .15$

$P[\theta_1^{ML} = \theta_1^{true}] \searrow$ with n

$P[|\theta_1^{ML} - \theta^{true}| < \varepsilon] \nearrow 1$ with n

$n = 200 \quad P[e < \varepsilon] \approx 1$

$P[e > \varepsilon | \theta_1^{ML}] \approx$

$\approx P[e > \varepsilon | \theta_1^{true}]$

$n = 20 \quad P[e < \varepsilon] \approx \frac{1}{2}$

Practically n large enough $\Rightarrow \theta_1^{ML} \approx \theta_1^{true} \Rightarrow$

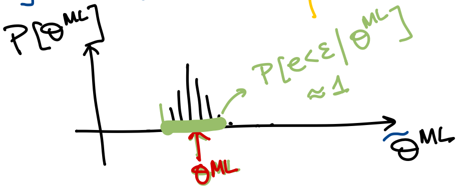
ML estimate as a random variable

In general n, Model class $= \{ P_\theta,\ \theta \in \text{some range} \}$

$n \rightarrow S_{\theta^{ML}} = \{$ all possible values $\}$

$P[\tilde{\theta}^{ML} \mid \theta^{true} = \theta^{ML}]$ on $S_{\theta^{ML}}$

Data $x^{1:n} \xrightarrow{ML} \theta^{ML}$



$P[e < \varepsilon \mid \theta^{ML}] \approx 1$

n was large enough

$\theta^{true} \approx \theta^{ML}$