STAT 391 4/20/23

# Lecture 8

Continuous distributions

uniform (continued)

# Lecture Notes IV – Continuous distributions. Parametric density estimation.

Marina Meilă
mmp@stat.washington.edu

Department of Statistics
University of Washington

April, 2023

CDF and PDF. Sampling

Examples of continuous distributions

ML estimation for continuous distributions

ML estimation by gradient ascent

**Reading**: Ch.5, 6

# CDF and PDF refresher

**Cumulative distribution function (CDF)**

$$F(x) = P[X \leq x] \qquad (1)$$

1. $F \geq 0$ positivity.
2. $\lim_{x \to -\infty} F = 0$
3. $\lim_{x \to \infty} F = 1$
4. $F$ is an increasing function

**Probability density [function] (PDF)**

$$f = \frac{dF}{dx} \qquad (2)$$

$$P(a, b) = P[a, b] = F(b) - F(a) = \int_a^b f(x)dx \qquad (3)$$

**normalization condition**

$$\int_{-\infty}^{\infty} f(x)dx = 1 \qquad (4)$$

# Examples of continuous distributions

Families

$$\mathcal{F}_1 = \{u_{[a,b]},\ a < b\} \quad \text{uniform} \tag{5}$$

$$f(x; a, b) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

$$\mathcal{F}_2 = \{N(.; \mu, \sigma^2)\} \quad \text{normal} \tag{7}$$

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{8}$$

$$F(x; a, b) = \frac{1}{1 + e^{-ax-b}}, \quad a > 0 \quad \text{logistic} \tag{9}$$

$$f(x; a, b) = \frac{ae^{-ax-b}}{(1 + e^{-ax-b})^2} \tag{10}$$

# ML estimation for continuous distributions = density estimation

$S = (-\infty, \infty)$ or $S \subset (-\infty, \infty)$

Sample space

Data $\mathcal{D} = \{x^1, x^2, \ldots x^n\} \subset S$ iid from true unknown $f$

Model family $\mathcal{F} = \{f(x|\theta)\}$, $\theta$ = parameter(s)

Normal, Uniform, ....

Chosen

Problem estimate $\theta$ from $\mathcal{D}$

Solution : Max Likelihood $\to \theta^{ML} \to f(x|\theta^{ML})$

estimated density

# ML estimation for continuous distributions

Likelihood of $\theta$

$F'(x$

$$L(\theta) = \prod_{i=1}^{n} f(x^i | \theta)$$

$P(x^i | \theta)$ for $S$ discrete

log-likelihood

$$\ell(\theta) = \sum_{i=1}^{n} \ln f(x^i | \theta)$$

Max Likelihood Principle

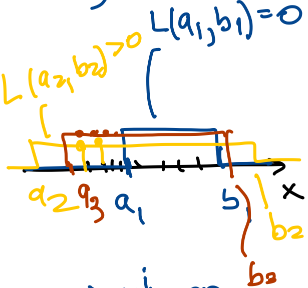
Choose $\theta^{ML} = \underset{\theta}{\operatorname{argmax}}\ \ell(\theta)$

# ML estimation for continuous distributions uniform

**Data** $\mathcal{D} = \{x^{1:n}\} \subset \underbrace{\mathbb{R}}_{S}$

**Model family**

$$\mathcal{F} = \{ u_{[a,b]},\ a<b,\ a,b\in\mathbb{R}\}$$




**Likelihood**

$$f(x) = \begin{cases} \frac{1}{b-a} & x\in[a,b] \\ 0 & \text{otherwise}\end{cases}$$

statistics

calculus

parameters: $\underline{a,b}$

$$L(a,b) = \prod_{i=1}^{n} f_{a,b}(x^i) = \begin{cases} 0 & \text{if } a > \min_{i=1:n} x^i \text{ or } b < \max_{i=1:n} x^i \\ \frac{1}{(b-a)^n} & \end{cases} \iff \min|b-a|$$

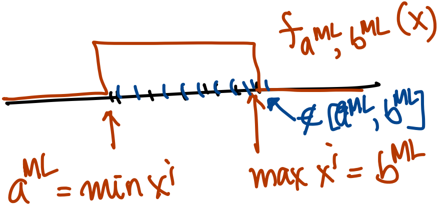
$\shortparallel$

$L(a,b|\mathcal{D})$

Maximize

# ML estimation for continuous distributions

$$\Rightarrow \boxed{\begin{aligned} a^{ML} &= \min_{i=1:n} x^i \\ b^{ML} &= \max_{i=1:n} x^i \end{aligned}}$$




Pb even if $x \sim \text{unif}[a^*, b^*]$ true distribution

$$\begin{matrix} a^{ML} > a^* \\ b^{ML} < b^* \end{matrix} \quad \Rightarrow \quad \begin{matrix} P[x \in [a^*, a)] > 0 \\ P[x \in (b, b^*)] > 0 \end{matrix}$$

but $f_{[a^{ML}, b^{ML}]}(x) = 0$ !! BAD!

# ML estimation for continuous distributions

$U_{[a,b]}$

**Solution** choose $\tilde{a} < a^{ML} = \min x^i$

$\tilde{b} > b^{ML} = \max x^i$

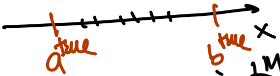
**Statistical solution 1**

(1) $1-\delta =$ confidence level

$\delta = Pr[\text{fail}]$

$[\tilde{a},\tilde{b}] \not\supset [a^{true}, b^{true}]$

(2) Assume $a^{true}, b^{true}$ known




$$Pr[\mathcal{D} \in [a^{ML}, b^{ML}] \mid a^{true}, b^{true}] = Pr[\max_{i=1:n} x^i = b^{ML}, \min x^i = a^{ML}]$$

$$l^* = b^{true} - a^{true}$$

$$l = b^{ML} - a^{ML} = \max \mathcal{D} - \min \mathcal{D}$$

# ML estimation for continuous distributions

Side problem $x^{1:n} \sim P$

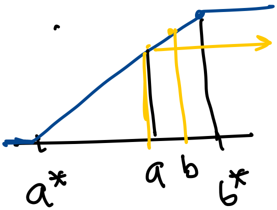
What is $\max x^{1:n} \sim P^{max}$ ?

$P[x < a] = F(a)$

$P[x^{1:n} < a] = \prod_{i=1}^{n} F(a) = F^n(a)$

iid ↑

for $U[a^*, b^*]$




$F(a) = \frac{a - a^*}{b^* - a^*}$ $\Rightarrow$ $F(b) - F(a) = \frac{b-a}{b^*-a^*}$

# ML estimation for continuous distributions

Side pb part 2:

$$Pr[x^{1:n} \in [a,b]] \underset{?}{=} (F(b)-F(a))^n$$

$$P[x \in [a,b]] = F(b) - F(a)$$

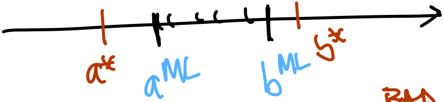
Back to uniform $[a^*, b^*]$ :

$$P[x^{1:n} \in [a,b] \mid a^*, b^*] = \left(\frac{b^{ML}-a^{ML}}{\underbrace{b^*-a^*}_{<1}}\right)^n$$

with $n$

$$= \left(\frac{\ell}{\ell^*}\right)^n = \delta$$

$Pr[\text{BAD}]$




BAD $\frac{\ell}{\ell^*}$ small

STATISTICS

calculus

# ML estimation for continuous distributions

$$\left(\frac{\ell}{\ell^*}\right)^n = \delta \Rightarrow \frac{\ell}{\ell^*} = \sqrt[n]{\delta} \Longrightarrow \boxed{\ell^* = \frac{\ell}{\sqrt[n]{\delta}}}$$

$b^{ML} - a^{ML}$ ← from data

confidence

$$\Rightarrow b^* - a^* = \ell^* = \boxed{(b^{ML} - a^{ML})\frac{1}{\sqrt[n]{\delta}}} = \Delta \quad \text{correction}$$

$$\ell^* - \ell = \ell\left(\delta^{-\frac{1}{n}} - 1\right)$$

$$a^* = a^{ML} - \frac{\ell^* - \ell}{2} = \boxed{a^{ML} - \frac{1}{2}\left(b^{ML} - a^{ML}\right)\left(\delta^{-\frac{1}{n}} - 1\right)}$$

$$b^* = b^{ML} + \frac{\ell^* - \ell}{2} = \boxed{b^{ML} - \frac{1}{2}\left(b^{ML} - a^{ML}\right)\left(\delta^{-\frac{1}{n}} - 1\right)}$$

→ corrected estimates

# ML estimation for continuous distributions

$$\boxed{\left(b^{ML} - a^{ML}\right) \frac{1}{\sqrt[n]{\delta}}} = \Delta \quad \text{correction}$$

$\delta \searrow$ want to be safer $\Rightarrow \frac{1}{\delta} \nearrow \Rightarrow \Delta \nearrow$

$n \nearrow$ more data safer

$\sqrt[n]{\delta} \nearrow 1$ for $n \to \infty \Rightarrow \Delta \searrow b^{ML} - a^{ML}$

$\delta \in (0,1)$

# ML estimation for continuous distributions

Exponential

$S = [0, \infty)$

$f_\lambda(x) = \lambda e^{-\lambda x}$ $\quad x > 0$, $\lambda > 0$

Model family $\mathcal{F} = \{ f_\lambda, \lambda > 0 \}$

Data $x^{1:n} > 0$

Likelihood $L(\lambda) = \prod_{i=1}^{n} \left( \lambda e^{-\lambda x^i} \right) = \lambda^n e^{-\lambda \sum_{i=1}^{n} x^i}$

log-$\ell$ $\quad \ell(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^{n} x^i$

STAT

CALC

ML

$$\lambda^{ML} = \underset{\lambda > 0}{\operatorname{argmax}}\, \ell(\lambda)$$

# ML estimation for continuous distributions

$$\ell(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^{n} x^i$$

$$\lambda^{ML} = \underset{\lambda > 0}{\operatorname{argmax}}\ \ell(\lambda)$$

$$\ell'(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} x^i = 0$$

$$\boxed{\frac{1}{\lambda^{ML}} = \boxed{\frac{1}{n}\sum_{i=1}^{n} x^i}} > 0$$

mean of data!

sufficient statistic

Uniform

Exponential Family

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