STAT 391

4/25/23

# Lecture 9

# 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

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

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

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

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

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

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

# ML estimation for continuous distributions

**Normal** $(\mu, \sigma^2)$

$S = \mathbb{R} \equiv (-\infty, \infty)$

$\mathcal{F} = \{ f_{\mu,\sigma^2},\ \underline{\mu} \in \mathbb{R},\ \underline{\sigma^2} > 0 \}$ → parameters

Data $\{x^{1:n}\} \subset (-\infty, \infty)$

Wanted ML estimates for parameters $\mu, \sigma^2$

(log) Likelihood

$$L(\mu, \sigma^2 \mid \mathcal{D}) = \prod_{i=1}^{n} f_{\mu,\sigma^2}(x^i) = \prod_{i=1}^{n} \frac{e^{-\frac{(x^i-\mu)^2}{2\sigma^2}}}{\sqrt{\sigma^2 2\pi}}$$

$$\max_{\mu,\sigma^2} \ell(\mu, \sigma^2) = \sum_{i=1}^{n} -\frac{(x^i-\mu)^2}{2\sigma^2} - \frac{1}{2}\ln(\sigma^2 2\pi)$$

STAT ⇄ CALCULUS

# ML estimation for continuous distributions

$$(\ln z)' = \frac{1}{z}$$

$$\ell(\mu,\sigma^2) = \sum_{i=1}^{n}\left[-\frac{(x^i-\mu)^2}{2\sigma^2} - \frac{1}{2}\ln(\sigma^2 2\pi)\right]$$

$$\frac{\partial \ell}{\partial \mu} = 0 = -\frac{1}{2\sigma^2}\frac{\partial}{\partial \mu}\sum_{i=1}^{n}(x^i-\mu)^2 = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}2(\mu - x^i) =$$

$$= -\frac{1}{\sigma^2}\left[n\mu - \sum_{i=1}^{n}x^i\right] \Rightarrow \boxed{\mu^{ML} = \frac{1}{n}\boxed{\sum_{i=1}^{n}x^i}}$$

sufficient statistics

$$\frac{\partial \ell}{\partial \sigma^2} = 0 = +\sum_{i=1}^{n}\frac{(x^i-\mu)^2}{2}\,\frac{+1}{(\sigma^2)^2} - \frac{n}{\sigma^2}\cdot\frac{1}{2} \Rightarrow$$

$$\Rightarrow \frac{1}{\sigma^2}\sum_{i}(x^i - \mu^{ML})^2 = n \Rightarrow \boxed{(\sigma^2)^{ML} = \frac{1}{n}\sum_{i=1}^{n}(x^i - \mu^{ML})^2}$$

# ML estimation for continuous distributions

$$\boxed{(\sigma^2)^{ML} = \frac{1}{n}\sum_{i=1}^{n}(x^i - \mu^{ML})^2} \rightarrow \text{sample variance !}$$

$$\sum_{i=1}^{n}(x^i - \mu^{ML})^2 = \sum_{i=1}^{n}(x^i)^2 - n(\mu^{ML})^2 \leftarrow \text{Exercise}$$

$$\boxed{(\sigma^2)^{ML} = \frac{1}{n}\boxed{\sum_{i=1}^{n}(x^i)^2} - (\mu^{ML})^2}$$

sufficient statistics

S.S.

- Only $\exists$ for exp family [& uniform]
- # params = # S.S.
- always sums or average

$$\tilde{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x^i - \mu^{ML})^2 = \frac{n}{n-1}(\sigma^2)^{ML}$$

CORRECTED

# ML estimation by gradient ascent

Logistic $S = (-\infty, \infty)$

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

$\mathcal{F} = \{ f_{ab},\ a > 0 \}$

Want ML estimators for $a, b$

$$f_{a,b}(x) = \frac{-a\, e^{-ax-b}}{(1+e^{-ax-b})^2}$$

Likelihood $L(a,b) = \prod_{i=1}^{n} f_{a,b}(x^i)$

Log. likelihood

$$\max_{a,b}\ l(a,b) = n \ln a - a \sum_i x_i - nb - 2 \sum_{i=1}^{n} \ln(1+e^{-ax_i - b})$$

$$\frac{\partial l}{\partial a} = \frac{n}{a} - \sum_{i=1}^{n} x_i \frac{1-e^{-ax_i-b}}{1+e^{-ax_i-b}} = 0$$

$$\frac{\partial l}{\partial b} = -\sum_{i=1}^{n} \frac{1-e^{-ax_i-b}}{1+e^{-ax_i-b}} = 0$$

No closed form solution !! :(

No sufficient statistics !! :(

Remark

$\exists$ unique $(a^{ML}, b^{ML}) = \operatorname{argmax}_{a,b} l$

# ML estimation by gradient ascent

Maximize $\ell(a,b)$ by gradient ascent

$$l(a,b) \;=\; n\ln a - a\sum_i x_i - nb - 2\sum_{i=1}^{n}\ln(1+e^{-ax_i-b})$$

$$\frac{\partial l}{\partial a} \;=\; \frac{n}{a} - \sum_{i=1}^{n} x_i\frac{1-e^{-ax_i-b}}{1+e^{-ax_i-b}} \;=\; 0$$

$$\frac{\partial l}{\partial b} \;=\; -\sum_{i=1}^{n}\frac{1-e^{-ax_i-b}}{1+e^{-ax_i-b}} \;=\; 0$$

**Algorithm**

Initialize $a^0, b^0$ $\quad a^0>0$

Until convergence $t=1,2,\ldots$

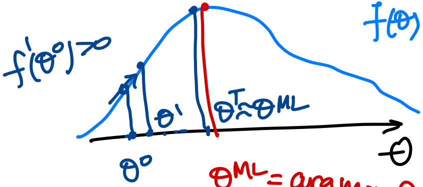
$$a^t \leftarrow a^{t-1} + \eta\frac{\partial \ell}{\partial a}(a^{t-1}, b^{t-1})$$

$$b^t \leftarrow b^{t-1} + \eta\frac{\partial \ell}{\partial b}(a^{t-1}, b^{t-1})$$

Output $a^t, b^t = a^{ML}, b^{ML}$

# ML estimation by gradient ascent

$\boxed{\eta = \text{step size}}$




$\theta^{ML} = \arg\max f$

$\theta^1 = \theta^0 + \eta \cdot f'(\theta^0)$

$\theta^T \approx \theta^{ML} \Rightarrow f'(\theta^T) \approx 0$

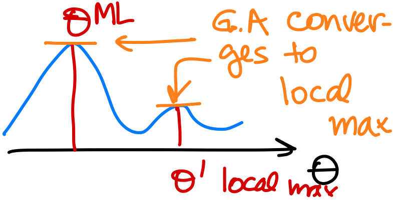
$\Downarrow$ STOP

"Convergence"

$$l(a,b) \;=\; n \ln a - a \sum_i x_i - nb - 2 \sum_{i=1}^n \ln(1 + e^{-ax_i - b})$$

$$\frac{\partial l}{\partial a} \;=\; \frac{n}{a} - \sum_{i=1}^n x_i \frac{1 - e^{-ax_i - b}}{1 + e^{-ax_i - b}} \;=\; 0$$

$$\frac{\partial l}{\partial b} \;=\; -\sum_{i=1}^n \frac{1 - e^{-ax_i - b}}{1 + e^{-ax_i - b}} \;=\; 0$$




Initialize

any $a > 0$, any $b$ OK

because no local maxima!