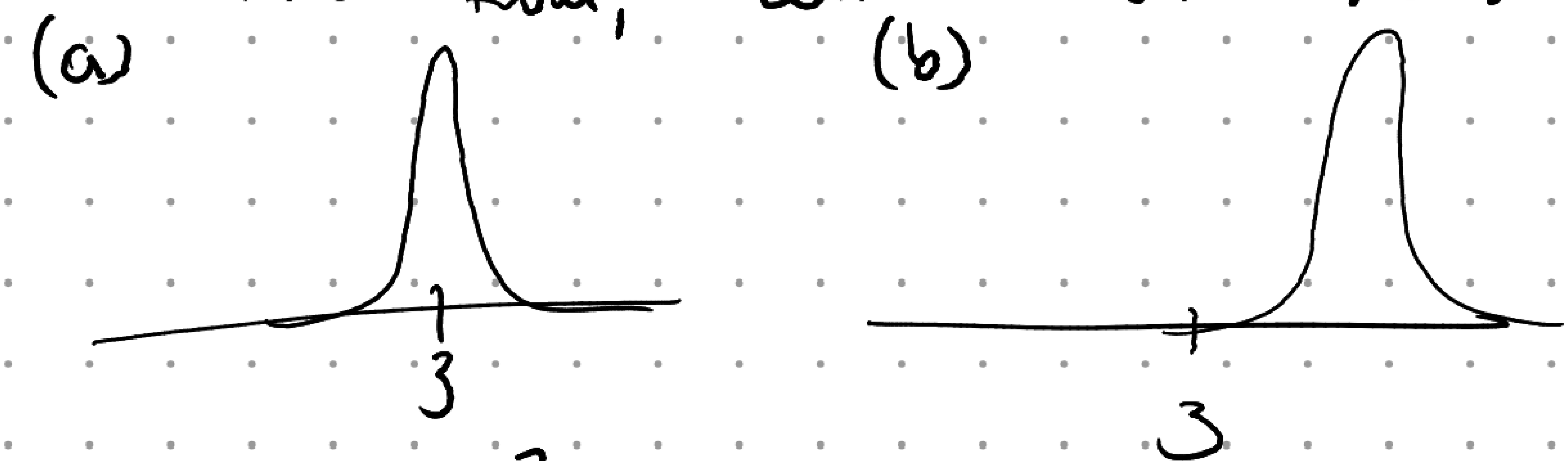
Maximum Likelihood

$\mathcal{F} = \{f_\theta : \theta \in \Theta\}$, $X_1, \dots, X_n \sim F_\theta$ for some fixed unknown θ.

How can we learn θ?

Ex. Suppose we observe $X_1 = 3$

If we had to guess the density f came from, which would be better?



$\mathbb{P}[X \in [3, 3+dx]] \approx f_\theta(3)\,dx$ this is a fxn of θ

Def. The likelihood function is defined by

$$\mathcal{L}_n(\theta) = \prod_{i=1}^{n} f_\theta(X_i)$$

Def. The log likelihood fnct is

$$\ell_n(\theta) = \log(\mathcal{L}_n(\theta)) = \sum_{i=1}^{n} \log(f_\theta(X_i))$$

Def. The maximum likelihood estimate $\hat{\theta}_n$ is the value of θ which maximizes $\mathcal{L}_n(\theta)$

Note: maximizer of $\mathcal{L}_n(\theta)$ and $\ell_n(\theta)$ are the same

maximizer of $\mathcal{L}_n(\theta)$ and $c\,\mathcal{L}_n(\theta)$ are the same (normalizing constants from density don't matter)

Ex. $X_1, \dots, X_n \sim \text{Ber}(p)$

$$f_p(x) = p^x (1-p)^{1-x} = \begin{cases} p & x = 1 \\ 1-p & x = 0 \end{cases}$$

$$\mathcal{L}_n(p) = \prod_{i=1}^{n} f_p(X_i) = p^{X_1 + \cdots + X_n} (1-p)^{n-(X_1 + \cdots + X_n)}$$

$$= p^S (1-p)^{n-S}$$

$S = X_1 + \cdots + X_n$

$\ell_n(p) = S \log(p) + (n-S) \log(1-p)$

$$\frac{d\,\ell_n(p)}{dp} = 0 \implies p = \frac{S}{n} \implies \hat{p} = \frac{S}{n} \text{ is MLE}$$

Ex. $X_1, \dots, X_n \sim \text{Unif}(0, \theta)$

$$f_\theta(x) = \begin{cases} 1/\theta & 0 \le x \le \theta \\ 0 & \text{o.w.} \end{cases}$$

$$\mathcal{L}_n(\theta) = \prod_{i=1}^{n} f_\theta(X_i)$$

$$= \begin{cases} 0 & \exists i \text{ st } X_i < 0 \text{ or } X_i > \theta \\ \frac{1}{\theta^n} & \text{o.w.} \end{cases}$$

$$= \begin{cases} 0 & \max\{X_1, \dots, X_n\} > \theta \\ \frac{1}{\theta^n} & \max\{X_1, \dots, X_n\} \le \theta \end{cases}$$