Chapter ML:IV (continued)

IV. Statistical Learning

- ❑ Probability Basics
- ❑ Bayes Classification
- ❑ MAP versus ML Hypotheses

MAP versus ML Hypotheses

The Frequentist Setting to Parameter Estimation

$$\theta \rightsquigarrow \begin{cases} D_1: \ \underset{\hat{\theta} \in \Theta}{\text{argmax}}\, P(D_1 \mid \hat{\theta}) \ \rightarrow \ \hat{\theta}_{\text{ML}_1} \end{cases}$$

MAP versus ML Hypotheses

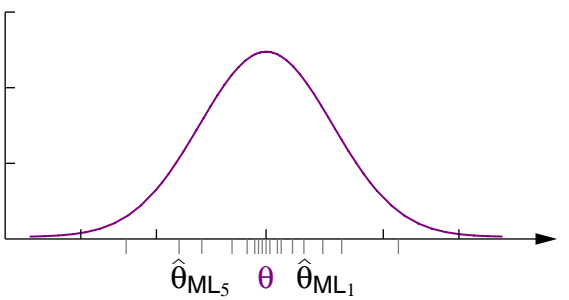
The Frequentist Setting to Parameter Estimation (continued)

$$\theta \rightsquigarrow \begin{cases} D_1: & \underset{\hat{\theta} \in \Theta}{\text{argmax}}\, P(D_1 \mid \hat{\theta}) & \rightarrow & \hat{\theta}_{\text{ML}_1} \\ \vdots & & & \vdots \\ D_k: & \underset{\hat{\theta} \in \Theta}{\text{argmax}}\, P(D_k \mid \hat{\theta}) & \rightarrow & \hat{\theta}_{\text{ML}_k} \end{cases}$$

MAP versus ML Hypotheses

The Frequentist Setting to Parameter Estimation (continued)

$$\theta \rightsquigarrow \begin{cases} D_1: & \underset{\hat{\theta}\in\Theta}{\text{argmax}}\, P(D_1 \mid \hat{\theta}) & \rightarrow & \hat{\theta}_{\text{ML}_1} \\ \vdots & & & \vdots \\ D_k: & \underset{\hat{\theta}\in\Theta}{\text{argmax}}\, P(D_k \mid \hat{\theta}) & \rightarrow & \hat{\theta}_{\text{ML}_k} \end{cases} \rightsquigarrow$$



Remarks:

- The sampling distribution quantifies $P(\hat{\theta} \mid \theta)$, the distribution of $\hat{\theta}$ given the real (but unknown) θ.
- For the sampling distribution holds: $P(\hat{\theta} \mid \theta) \propto N\left(\theta, \frac{\sigma^2}{n}\right)$, where n is the sample size of the D_i. This result is independent of the distribution family considered for the mechanism that generates D.