Bayesian approach to statistics

Different approaches to statistics

Frequentist

Bayesian

Uncertainty interpretation

Frequentist

Bayesian

Subjective

Objective

Data and parameters

Frequentist

θ is random
X is fixed

θ is fixed
X is random

Bayesian

Data and parameters

Frequentist

For any $|X|$

$|X| \gg |\theta|$

Bayesian

Training

Frequentist

Bayesian

Maximum Likelihood:

$$\widehat{\theta} = arg\max_{\theta} P(X|\theta)$$

Training

Frequentist

Bayesian

Bayes theorem:

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)}$$

Classification

Training:

$$P(\theta|X_{\text{tr}}, y_{\text{tr}}) = \frac{P(y_{\text{tr}}|X_{\text{tr}}, \theta)P(\theta)}{P(y_{\text{tr}}|X_{\text{tr}})}$$

Prediction:

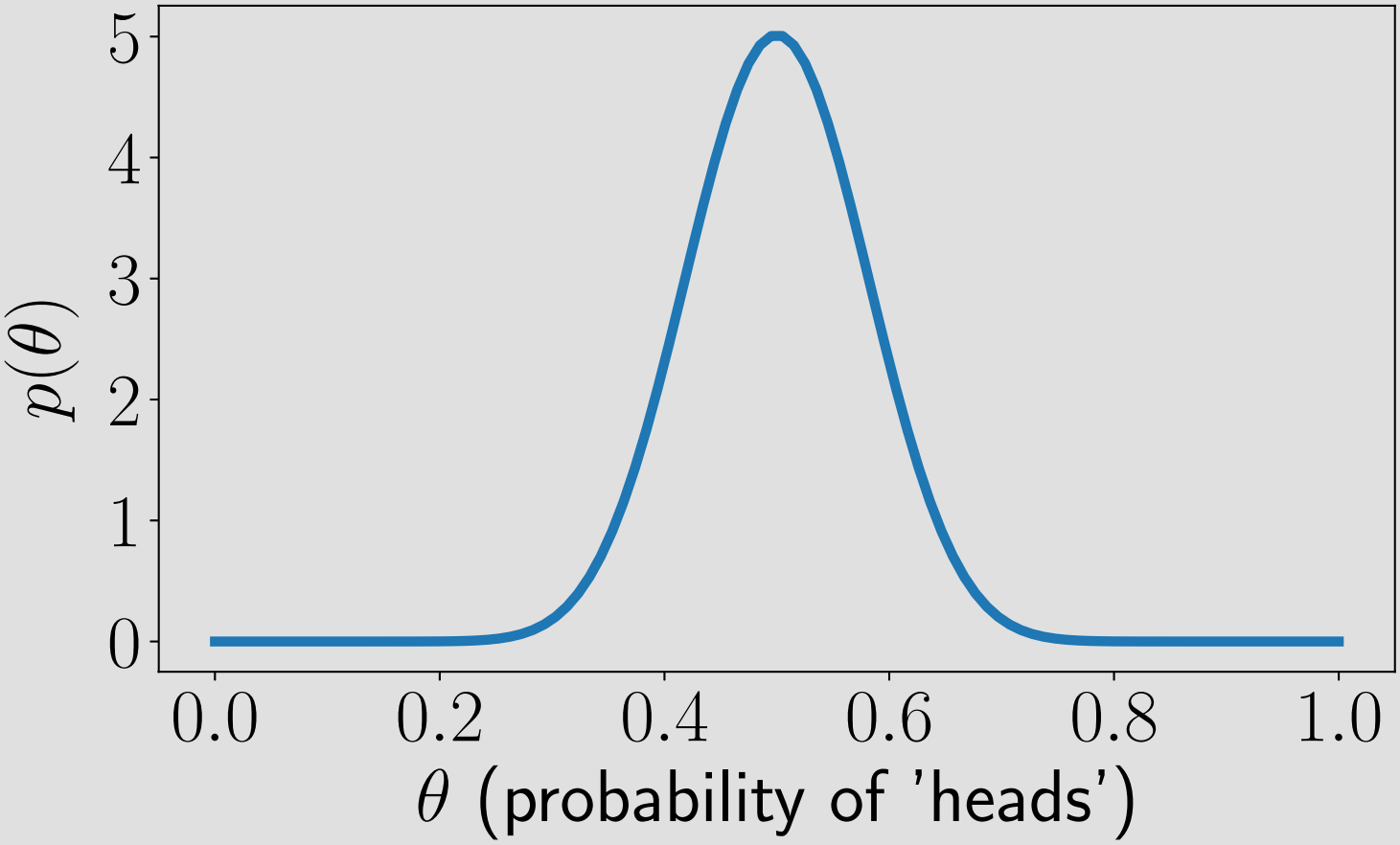
$$P(y_{\text{ts}}|X_{\text{ts}}, X_{\text{tr}}, y_{\text{tr}}) = \int P(y_{\text{ts}}|X_{\text{ts}}, \theta)P(\theta|X_{\text{tr}}, y_{\text{tr}})d\theta$$

Regularization

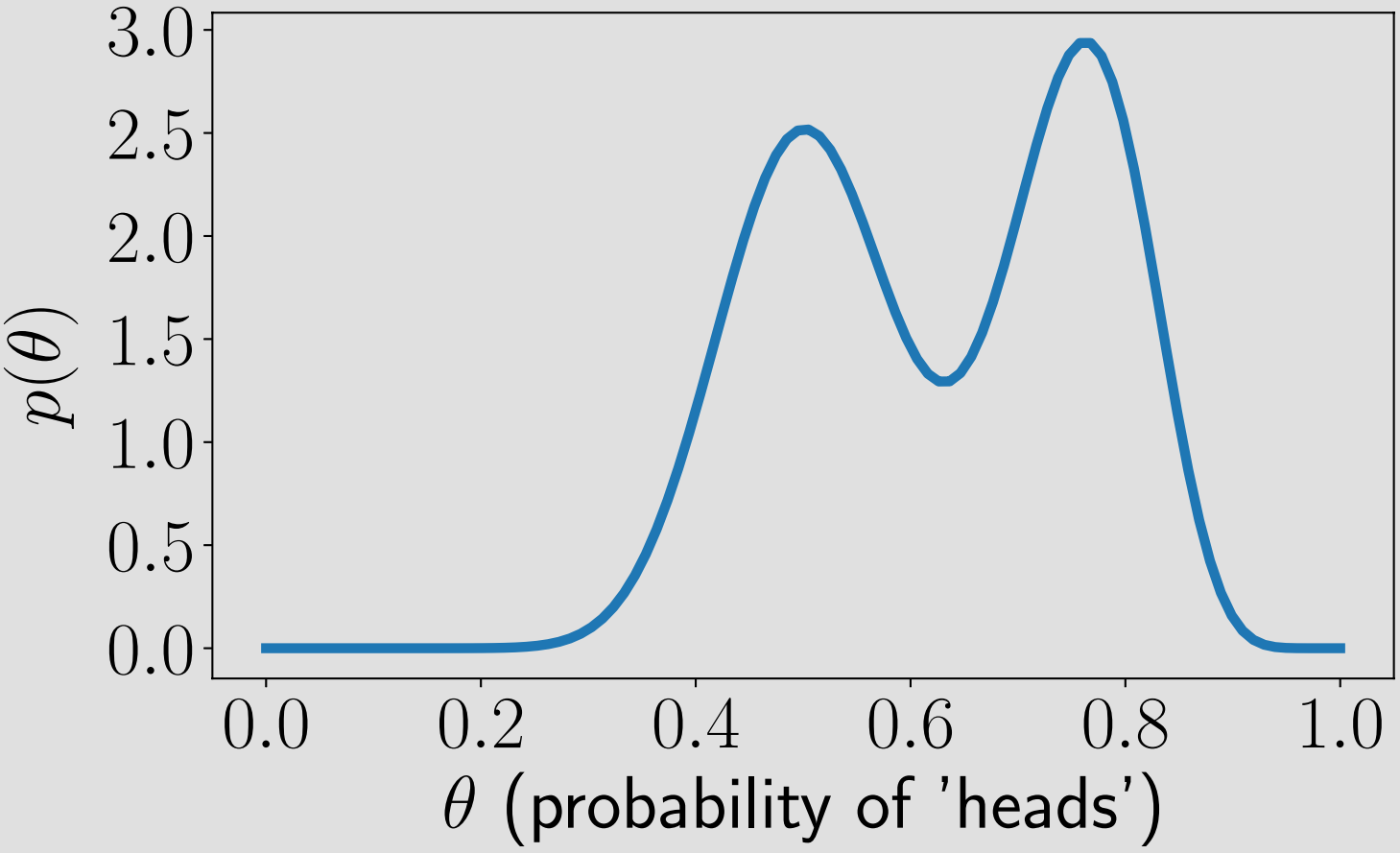
Regularizer

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)}$$

Regularization



Regularization



On-line learning

New prior

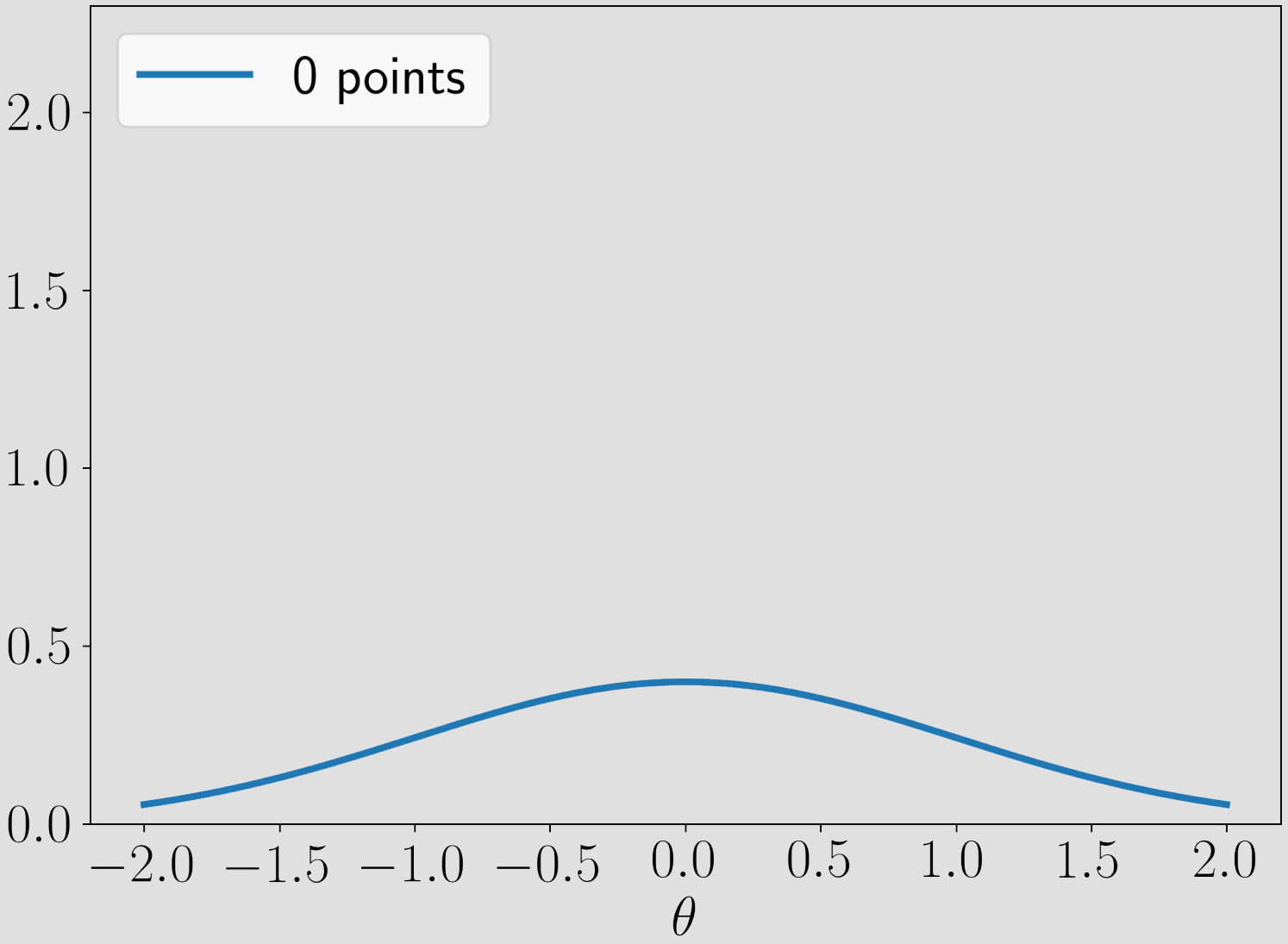
Likelihood

Prior

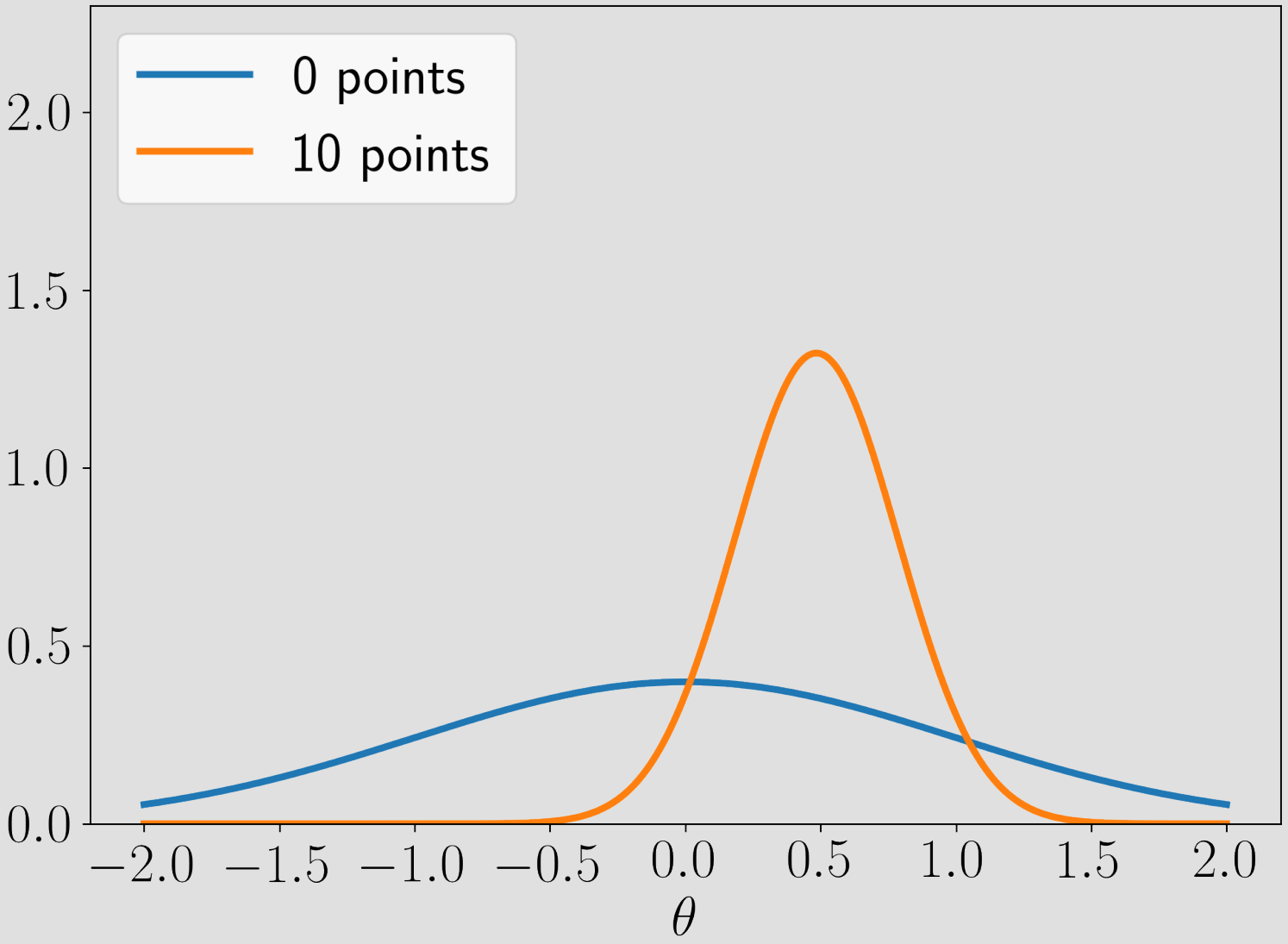
$$P_k(\theta) = P(\theta|x_k) = \frac{P(x_k|\theta)P_{k-1}(\theta)}{P(x_k)}$$

Posterior

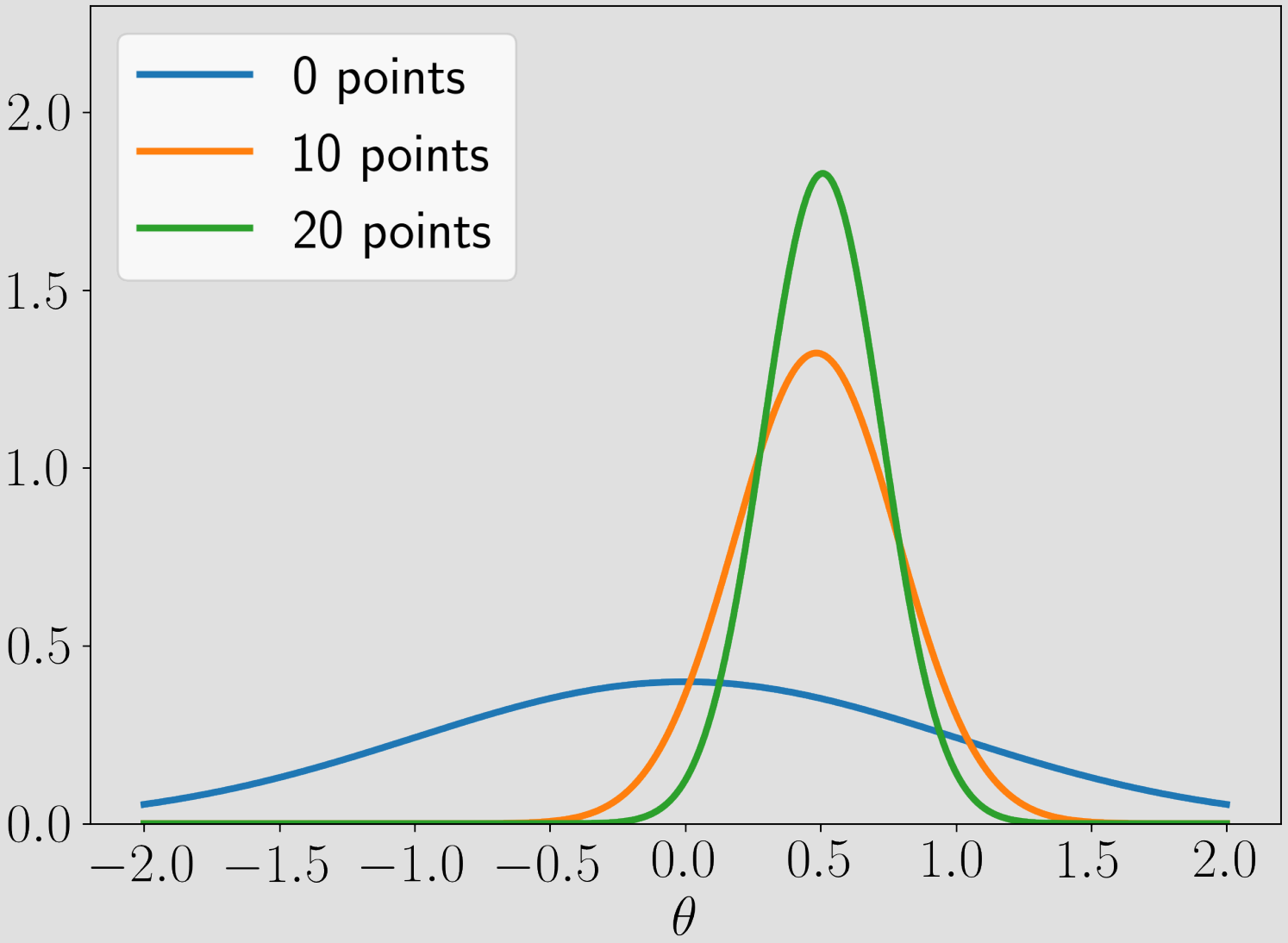
On-line learning



On-line learning



On-line learning



On-line learning