

2. Conjugate distributions



Bayes formula

Fixed by model Our own choice!

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

Fixed by data



Conjugate prior

$P(\theta)$ is **conjugate** to $P(X|\theta)$:

$$\mathcal{A}(v') \rightarrow P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)} \leftarrow \mathcal{A}(v)$$



Example

$$P(X|\theta) = \mathcal{N}(X|\theta, \sigma^2)$$

$$\mathcal{A}(v) = ?$$

$$\mathcal{A}(v') \rightarrow P(\theta|X) = \frac{\mathcal{N}(X|\theta, \sigma^2) P(\theta)}{P(X)}$$

Diagram illustrating the components of the posterior probability formula $P(\theta|X)$:

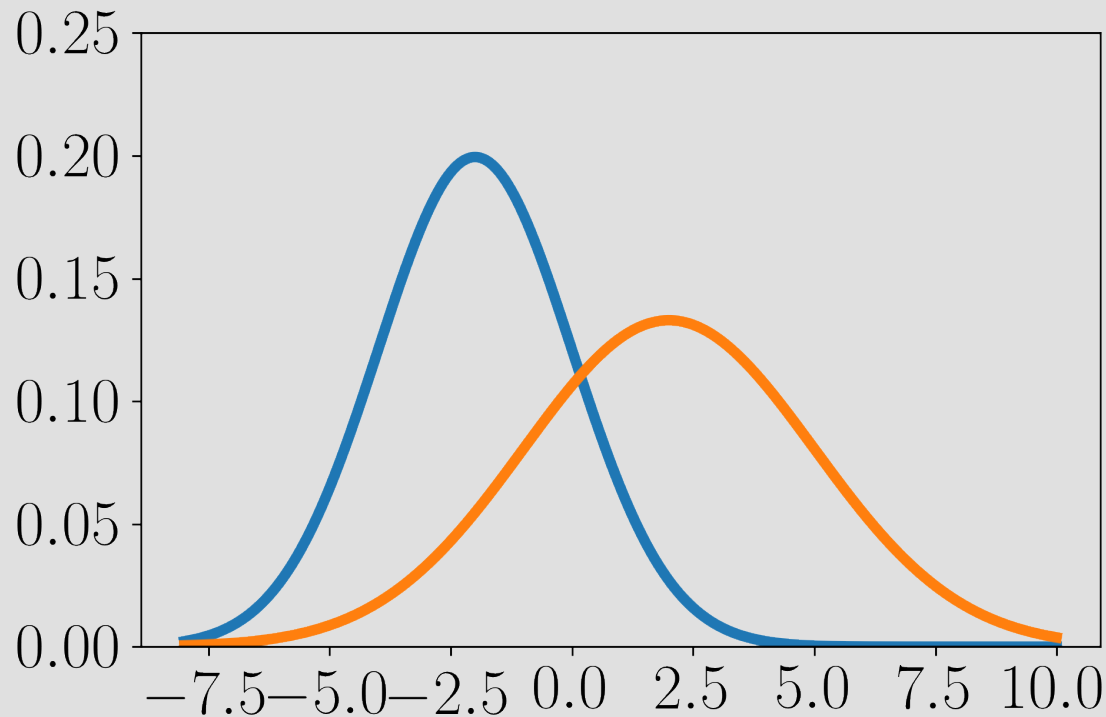
- $\mathcal{A}(v')$ points to the posterior $P(\theta|X)$.
- $\mathcal{N}(X|\theta, \sigma^2)$ points to the likelihood term in the numerator.
- $\mathcal{A}(v)$ points to the prior $P(\theta)$ in the numerator.



Two Gaussians

$$P(X_1) \sim \mathcal{N}(\mu_1, \sigma_1^2) \quad P(X_2) \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

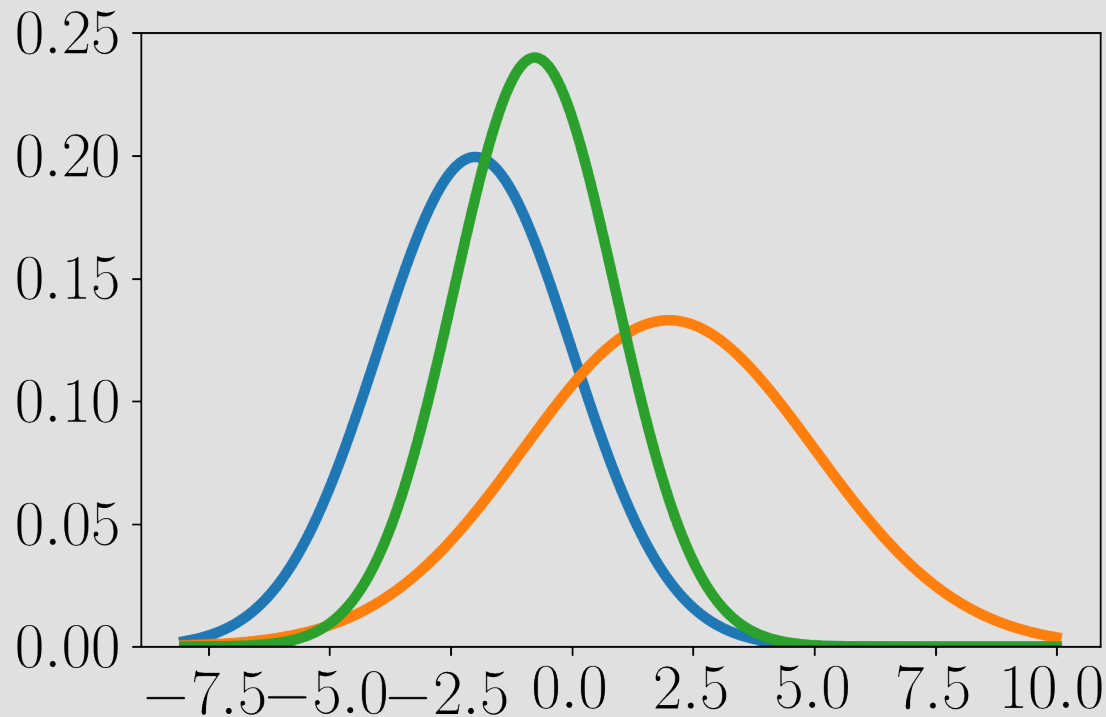
$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \text{const} \cdot e^{-\text{parabola}}$$



Two Gaussians

$$P(X_1) \sim \mathcal{N}(\mu_1, \sigma_1^2) \quad P(X_2) \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \text{const} \cdot e^{-\text{parabola}}$$



Solution

$$P(X|\theta) = \mathcal{N}(X|\theta, \sigma^2)$$

$$\mathcal{A}(v) = \mathcal{N}(\theta|a, b^2)$$

$$\begin{array}{ccc} \mathcal{N}(X|\theta, \sigma^2) & & \mathcal{N}(\theta|m, s^2) \\ \downarrow & & \downarrow \\ P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)} \\ \uparrow & & \\ \mathcal{N}(\theta|a, b^2) & & \end{array}$$



Example

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{\mathcal{N}(x|\theta, 1)\mathcal{N}(\theta|0, 1)}{p(x)}$$



Example

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{\mathcal{N}(x|\theta, 1)\mathcal{N}(\theta|0, 1)}{p(x)}$$

$$p(\theta|x) \propto e^{-\frac{1}{2}(x-\theta)^2} e^{-\frac{1}{2}\theta^2}$$



Example

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{\mathcal{N}(x|\theta, 1)\mathcal{N}(\theta|0, 1)}{p(x)}$$

$$p(\theta|x) \propto e^{-\frac{1}{2}(x-\theta)^2} e^{-\frac{1}{2}\theta^2}$$

$$p(\theta|x) \propto e^{-(\theta - \frac{x}{2})^2}$$



Example

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{\mathcal{N}(x|\theta, 1)\mathcal{N}(\theta|0, 1)}{p(x)}$$

$$p(\theta|x) \propto e^{-\frac{1}{2}(x-\theta)^2} e^{-\frac{1}{2}\theta^2}$$

$$p(\theta|x) \propto e^{-(\theta - \frac{x}{2})^2}$$

$$p(\theta|x) = \mathcal{N}(\theta|\frac{x}{2}, \frac{1}{2})$$

