

Model selection

M_1, M_2, \dots, M_K

D

$$p(M_i | D) \propto p(M_i) p(D | M_i)$$

marginal likelihood

$$p(x|D) \approx \frac{1}{K} \sum_{r=1}^K p(x|\theta_r) p(\theta_r|D)$$

$$p(\theta | D, M) = \frac{p(\theta | M) p(D | \theta, M)}{p(D | M)}$$

BIC, AIC, TIC

$$\log p(D | M) \approx \text{const} + \log p(D | \theta_M) - \text{penalty}$$

$$KL(p_{\text{data}} \parallel p_{\text{model}}) \rightarrow \min \Leftrightarrow \log p_{\text{model}}(D | \theta) \rightarrow \max_{\theta}$$

$$\frac{L(\theta; D)}{D}$$