Occam's razor & Bayesian model selection

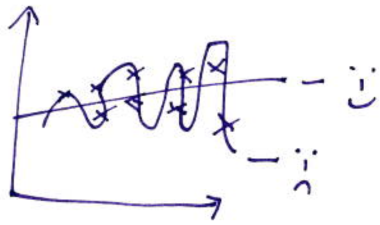
Problem $x \sim P(x)$ on $[-1,1]$

$H_0 : P(x|H_0) = \mathrm{Unif}([-1,1])$
$= \frac{1}{2} \mathbb{1}_{[-1,1]}$

$H_1 : P(x|H_1) = \frac{1}{2}(1+mx)$

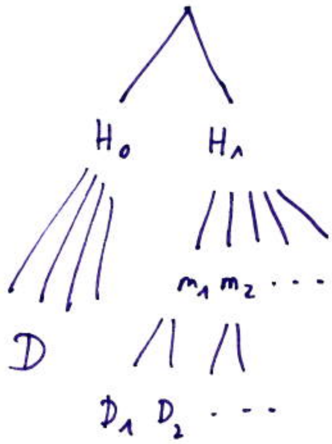
$m \in [-1,1]$

D data



find $\mathbb{P}(H_i|D)$

$\mathbb{P}(H_0) = \cancel{\tfrac{1}{2}} = \mathbb{P}(H_1)$ priors on hypothesis



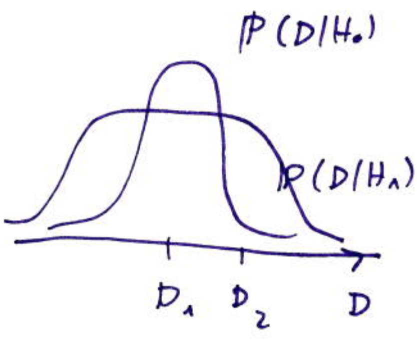
stupid method

$$\mathbb{P}(H_0|D) = \frac{\mathbb{P}(D|H_1)\cdot\mathbb{P}(H_0)}{\mathrm{const}\,(\mathbb{P}(D))}$$

$$\mathbb{P}(H_1|D) = \frac{\boxed{\mathbb{P}(D|H_1)}\,\mathbb{P}(H_1)}{\mathbb{P}(D)}$$

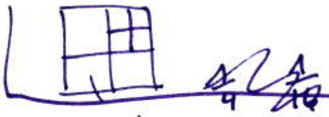
$$\mathbb{P}(D|H_1) = \int \mathbb{P}(D|H_1,m)\,\mathbb{P}(m|H_1)\,dm$$

$\cancel{\mathbb{P}(H_1)}$



$$\frac{\mathbb{P}(H_1|D)}{\mathbb{P}(H_0|D)} = \underbrace{\frac{\mathbb{P}(D|H_1)}{\mathbb{P}(D|H_0)}}_{\text{inc. Occam's razor}} \cdot \frac{\mathbb{P}(H_1)}{\mathbb{P}(H_0)}$$

Bayesian model selection:



$$\sum_{n=1}^{\infty}\left(\frac{1}{4}\right)^n = 1/3$$

1. For each model H_i find the optimal parameters with error bars (or better: find the distribution $P(\theta_i \mid D, H_i)$)

m

$$c = -\frac{\partial^2}{\partial x^2} L(x)\Big|_{x=x^*}$$

2. $\mathbb{P}(H_i \mid D) \sim \mathbb{P}(D \mid H_i)\, \mathbb{P}(H_i)$

$$= \int \mathbb{P}(D \mid H_i, \theta_i) \cdot \mathbb{P}(\theta_i \mid H_i)\, d\theta_i \cdot \mathbb{P}(H_i)$$

$$\mathbb{P}(H_0) = \frac{1}{2} = \mathbb{P}(H_1)$$

$$P(x \mid H_0) = \frac{1}{2} \cdot \mathbb{1}_{[-1,1]} \qquad \mathbb{P}(m \mid H_1) = \frac{1}{2} \qquad m \in [-1, 1]$$

$$P(x \mid H_1, m) = \frac{1}{2}(1 + mx)$$

$$D = \{x_1, x_2, \ldots, x_N\}$$

$$P(D \mid H_0) = \left(\frac{1}{2}\right)^N$$

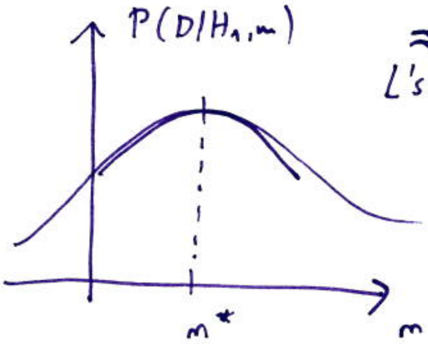
$$P(D \mid H_1, m) = \prod_{i=1}^{N} \frac{1}{2}(1 + m x_i)$$

$\rightsquigarrow$ „best" $m = m^*$

$$\mathbb{P}(D \mid H_1, m^*) = \max \mathbb{P}(D \mid H_1, m)$$

$$\cancel{\frac{P(D \mid H_1, m)}{P(D \mid H_0)} = \prod}$$

$$\mathbb{P}(D \mid H_1) = \int \overbrace{\mathbb{P}(D \mid H_1, m) \cdot \mathbb{P}(m \mid H_1)}^{\text{a posteriori } F(m)}\, dm$$

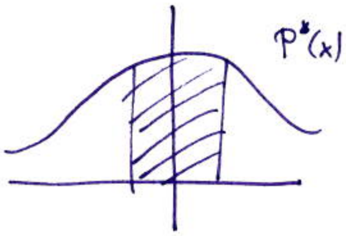


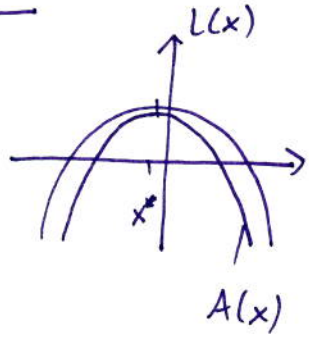
$$\underset{L\text{'s method}}{\approx} \mathbb{P}(D \mid H_1, \tilde{m}^*)\, \mathbb{P}(\tilde{m}^* \mid H_1) \cdot \frac{\sqrt{2\pi}}{-\frac{\partial^2}{\partial x^2} \log F(m) \Big|_{m=\tilde{m}}}$$

Laplace's method:

Given an unnormalized density $P^*(x)$, what ist $\int P^*(x)\,dx$?

Ex: $p^* = e^{-x^2}$





① $L(x) := \log P^*(x)$

② find $\operatorname{argmax} L(x) = x^*$

$$\Rightarrow L(x) \approx L(x^*) - \frac{c}{2}(x-x^*)^2 = A(x)$$

③ $\Rightarrow P^*(x) \approx \exp(A(x))$

$$= P^*(x^*)\, e^{-\frac{c}{2}(x-x^*)^2} = P^*(x^*)\, \frac{\sqrt{2\pi}}{\sqrt{c}} \underbrace{\frac{1}{\sqrt{2\pi}\frac{1}{\sqrt{c}}}\, e^{-\frac{c}{2}(x-x^*)^2}}_{\mathcal{N}(x^*,\, \sigma^2 = \frac{1}{c})}$$

$$\Rightarrow \int P^*(x)\,dx \approx P^*(x^*) \cdot \frac{\sqrt{2\pi}}{\sqrt{c}}$$

:-)

$P(D \mid H_0) = \left(\frac{1}{2}\right)^N$