BASICS OF BAYESIAN STATISTICS

The Bayes theorem states that

$$p(b|a) = \frac{p(a|b) \cdot p(b)}{p(a)}$$

PROOF Using the laws of probability we have

$$p(a, b) = p(a|b)\,p(b) = p(b|a)\,p(a)$$

$$\Rightarrow \quad p(b|a) = \frac{p(a|b)\,p(b)}{p(a)}$$

NOTE $p(a)$ is called marginal likelihood or evidence and can be written as

$$p(a) = \int p(a|b)\,p(b)\,db$$

EXAMPLE

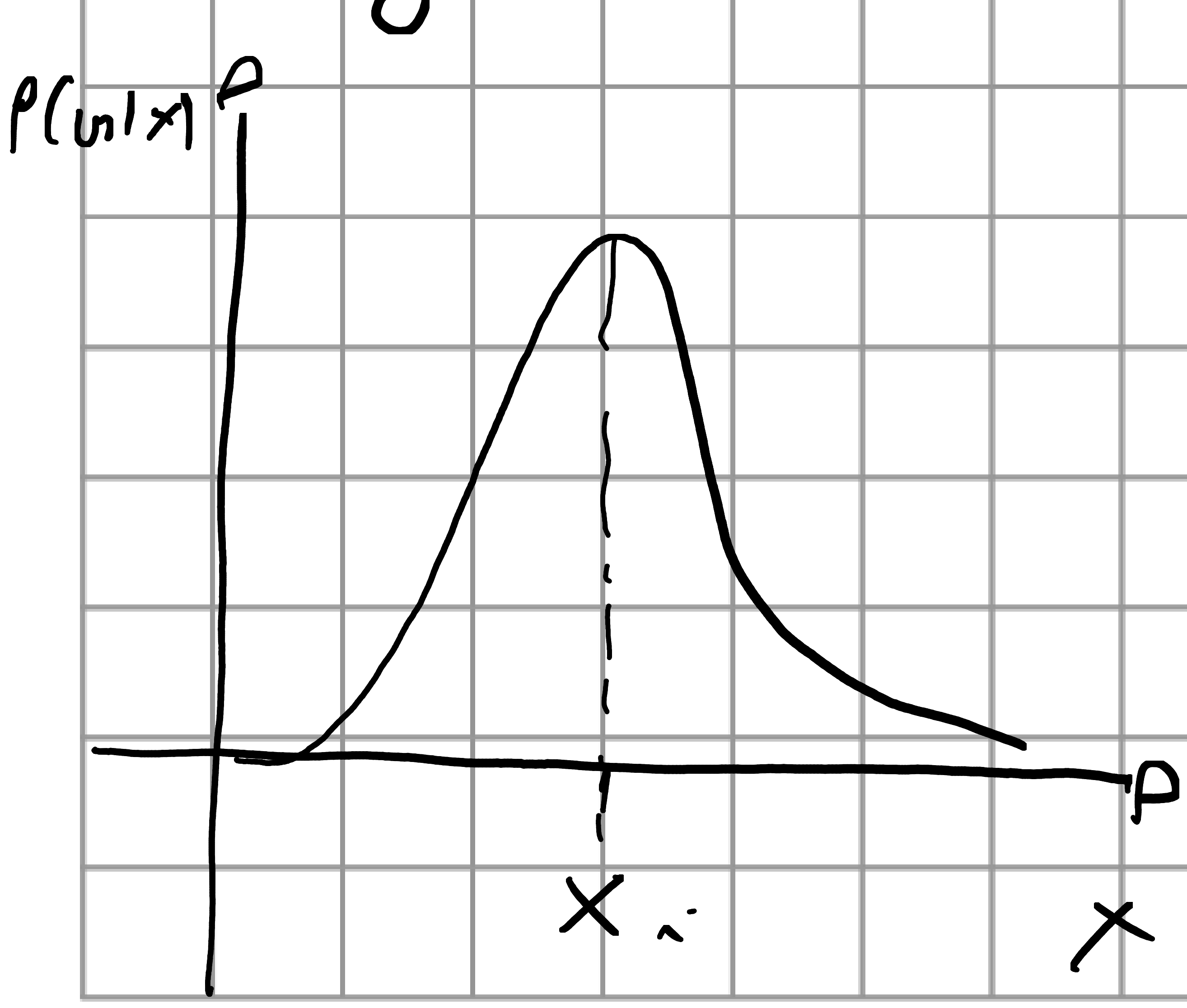
We want to estimate the mean of a Gaussian distribution. The first step is to write the our likelihood

$$\mathcal{L}(x \mid \mu) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

If we observe one value of x_i, then we have that the posterior is

$$P(\mu \mid x_i) = \frac{\mathcal{L}(x_i \mid \mu)\, P(\mu)}{\int \mathcal{L}(x_i \mid \mu)\, P(\mu)\, d\mu} = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x_i-\mu)}{2\sigma^2}}$$

We can observe that our posterior is a Gaussian with $\mu = x_i$ and $\sigma^2 = \sigma^2$



$$\langle \mu \rangle = \int \mu\, P(\mu \mid x_i)\, d\mu = \mu$$

$$\langle \sigma_\mu^2 \rangle = \int P(\mu \mid x_i)\, (\mu - \langle \mu \rangle)^2\, d\mu = \sigma^2$$

What happens when we have multiple observations? $\{X\} = \{x_1, \dots, x_n\}$

If the observations are I.I.D., then

$$\mathcal{L}(\{x\}|\mu) = \prod_{i=1}^{N} \mathcal{L}(x_i|\mu)$$

$$= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x_i-\mu)}{2\sigma^2}} = \left[\frac{1}{\sqrt{2\pi}\,\sigma}\right]^N \exp\left[\sum_{i=1}^{N} \frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

BUT $$\sum_{i=1}^{N} \frac{(x_i-\mu)^2}{2\sigma^2} = \sum \frac{(x_i^2 - 2x_i\mu + \mu^2)}{2\sigma^2} = -\frac{(\bar{x}-\mu)^2}{2\bar{\sigma}^2}$$

with $\bar{x} = \sum_i \frac{x_i}{N}$ and $\bar{\sigma}^2 = \frac{\sigma^2}{N}$

IMPORTANT It follows that the overall posterior will scale as $\sqrt{N}$ in terms of precision.

UNDERSTANDING SOME BASIC CONCEPTS

THE CONCEPT OF MODEL SELECTION

In Bayesian statistic we can also compare models. We can define the preference among competing models M_1 and M_2 as

$$\frac{P(M_1 | X)}{P(M_2 | X)} = \frac{P(X | M_1)}{P(X | M_2)} \boxed{\frac{P(M_1)}{P(M_2)}} \rightarrow \text{ODDS RATIO}$$

$$P(X | M_1) = \int P(X | \theta, M_1) \cdot P(\theta | M_1) \, d\theta$$

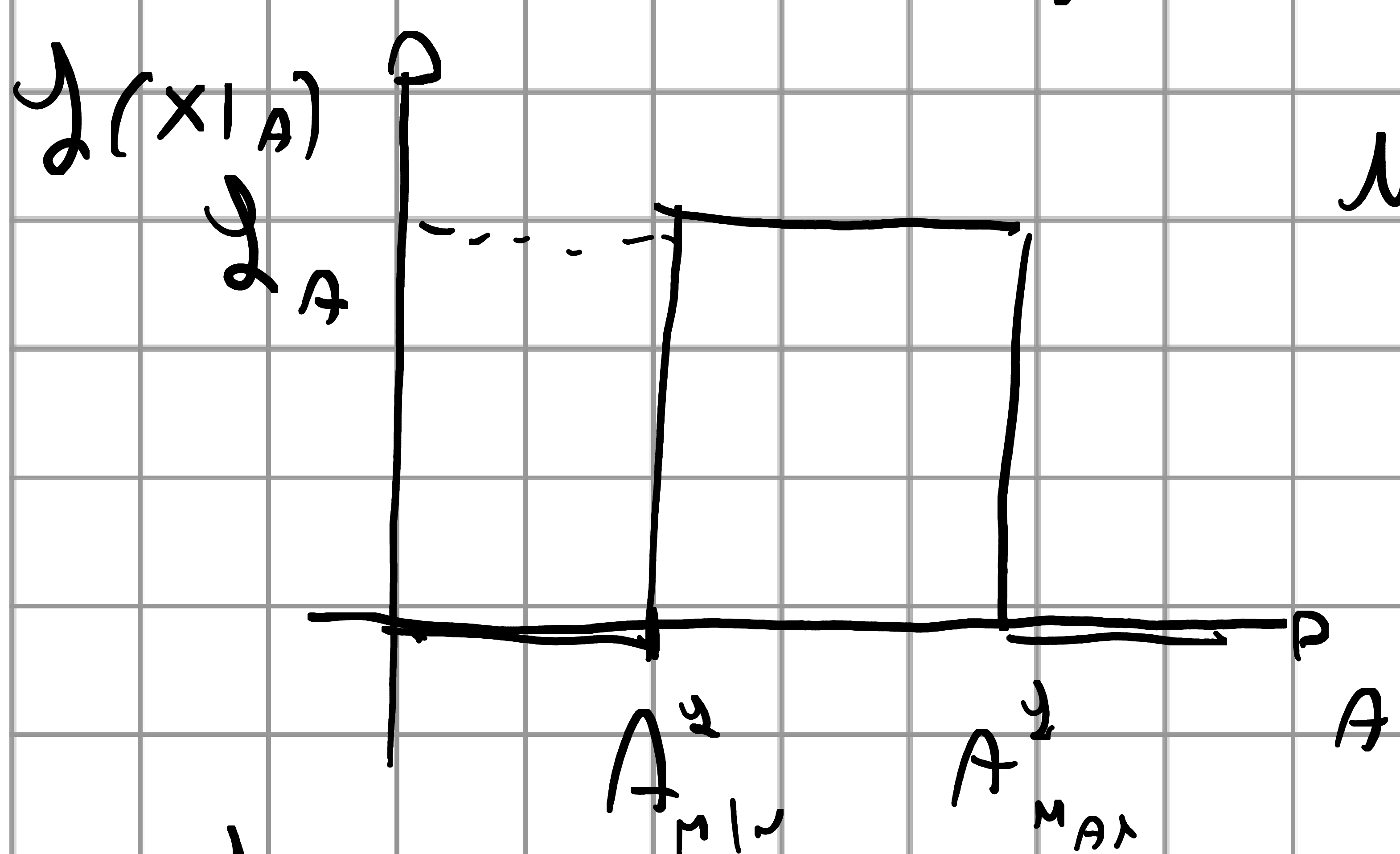
$\hookrightarrow$ MARGINAL LIKELIHOOD.

The two models can also have different # parameters.

EXAMPLE Imagine that we want to compare

The preference among two models. M_A, M_B

M_A: Has uniform likelihood to explain the distribution X, with a parameter A and has a prior $\pi(A) = \frac{\theta(|A_{Max}-A_{min}|)}{A_{Max}-A_{Min}}$



$$\mathcal{L}_A = \int_{A_{MIN}}^{A_{MAX}} \mathcal{L}(X|A) \frac{1}{A_{MAX}-A_{MIN}} \, dA$$

$$= \mathcal{L}_A \frac{A^1_M - A^1_m}{A_{MAX} - A_{MIN}}$$

. We see that the evidence is the maximum likelihood times posterior volume over prior volume.

PP-PLOT

$$P(a|b) = \frac{P(b|a)\,P(b)}{P(a)} \qquad C = \int_{-\infty}^{a} P(a|b_T)\, da$$

$$\frac{\partial C}{\partial a}$$