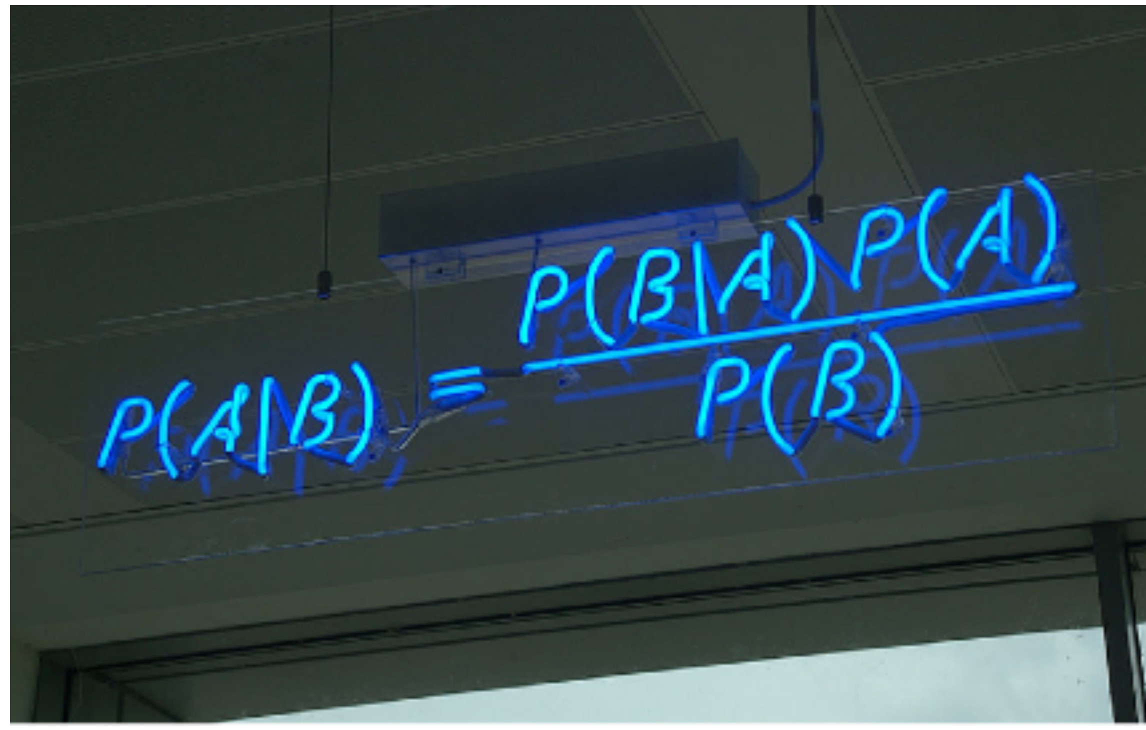
Bayesian Inference

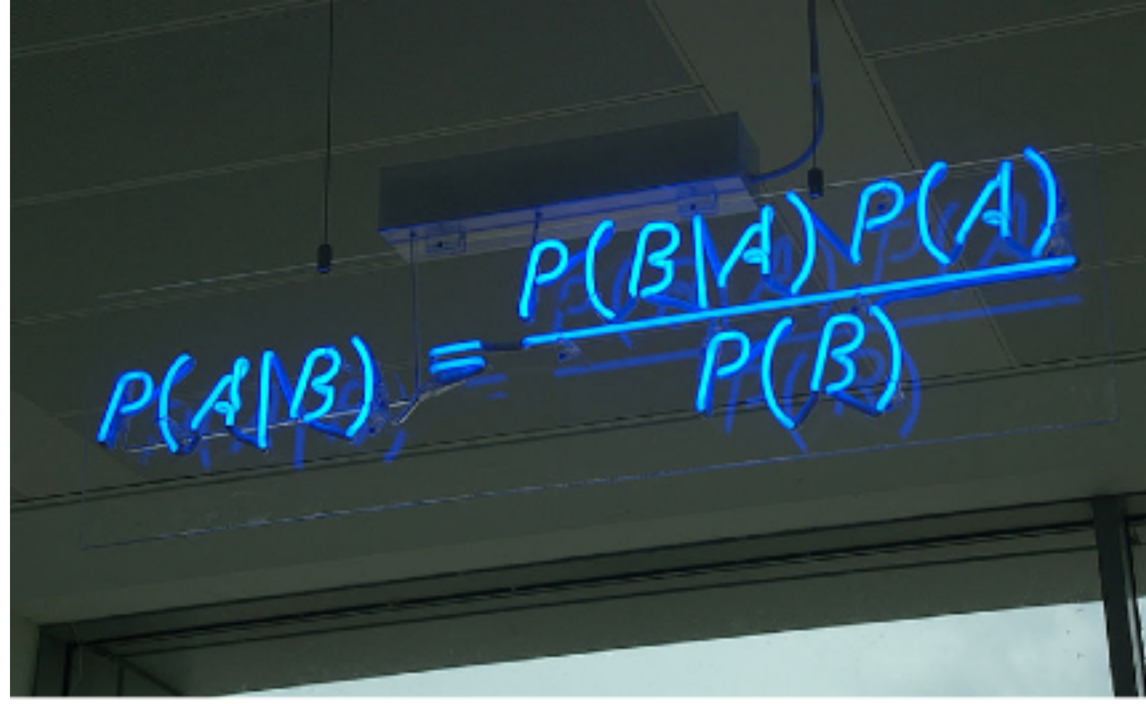


- Suppose you have data sampled from a parametric distribution with parameter θ.
- Our goal is to estimate θ, given the data samples.

$$\mathbb{P}(\theta\,|\,\text{data}) = \frac{\mathbb{P}(\text{data}\,|\,\theta) \times \mathbb{P}(\theta)}{\mathbb{P}(\text{data})}$$

- Posterior distribution: Represents what you know after having seen the data.
- Likelihood: How likely is the data generated from the distribution with parameter θ
- Prior: Represents what you know before seeing the data.
- $\mathbb{P}(\text{data}) = \int \mathbb{P}(\text{data}\,|\,\theta)\mathbb{P}(\theta)d\theta$: This is a normalisation factor, which is independent of θ. In general, this quantity is very difficult to compute.

Bayesian Inference



- Suppose you have data sampled from a parametric distribution with parameter θ.

- Our goal is to estimate θ, given the data samples.

$$\mathbb{P}(\theta \,|\, \text{data}) = \frac{\mathbb{P}(\text{data} \,|\, \theta) \times \mathbb{P}(\theta)}{\mathbb{P}(\text{data})}$$

- Let the data samples be $x_1, x_2, \ldots, x_n$.

- Then the posterior update given a new data sample x_{n+1} is

$$\mathbb{P}(\theta \,|\, x_1, \ldots, x_n, x_{n+1}) = \frac{\mathbb{P}(x_{n+1} \,|\, \theta)\mathbb{P}(\theta \,|\, x_1, \ldots, x_n))}{\sum_{\theta \in \Theta} \mathbb{P}(\theta \,|\, x_1, \ldots, x_n)\mathbb{P}(x_{n+1} \,|\, \theta)}$$