# Summary of Probabilistic Inference

## Everything follows from two simple rules

sum rule : $$p(x) = \overbrace{\sum_y}^{\text{or integral}} p(x,y)$$

product rule : $$p(x,y) = p(x)\,p(y|x)$$

## Learning

$$\overbrace{p(\theta|D,m)}^{\text{what we know about parameters after seeing data}} = \frac{\overbrace{p(D|\theta,m)}^{\text{what data tell us}}\;\overbrace{p(\theta|m)}^{\text{what we knew before}}}{p(D|m)}$$

data, model parameters, model

$$p(\overbrace{D}^{\text{data}}|\overbrace{\theta}^{\text{model parameters}},\overbrace{m}^{\text{model}}) = \text{likelihood of } \theta \text{ in model } m$$

$$p(\theta|m) = \text{prior probability of } \theta$$

## Prediction

$$p(x^*|D,m) = \int p(x|\theta,D,m)\,p(\theta|D,m)\,d\theta$$

$$\overset{\text{often}}{=} \int p(x|\theta,m)\,p(\theta|D,m)\,d\theta$$

average all possible predictions weighted by posterior probability

## Model Comparison

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

# Bayesian Decision Theory

expected conditional reward

$$\underbrace{R(a)}_{\text{action}} = \underbrace{\sum_x}_{\text{sum over all possible world states}} \overbrace{R(a,x)}^{\text{reward for carrying out action } a \text{ in world } x}\;\underbrace{p(x|D)}_{\text{posterior probability of world state } x \text{ given data } D}$$

$\Rightarrow$ compute action with highest expected conditional reward

$\Rightarrow$ SEPARATES INFERENCE & DECISION MAKING