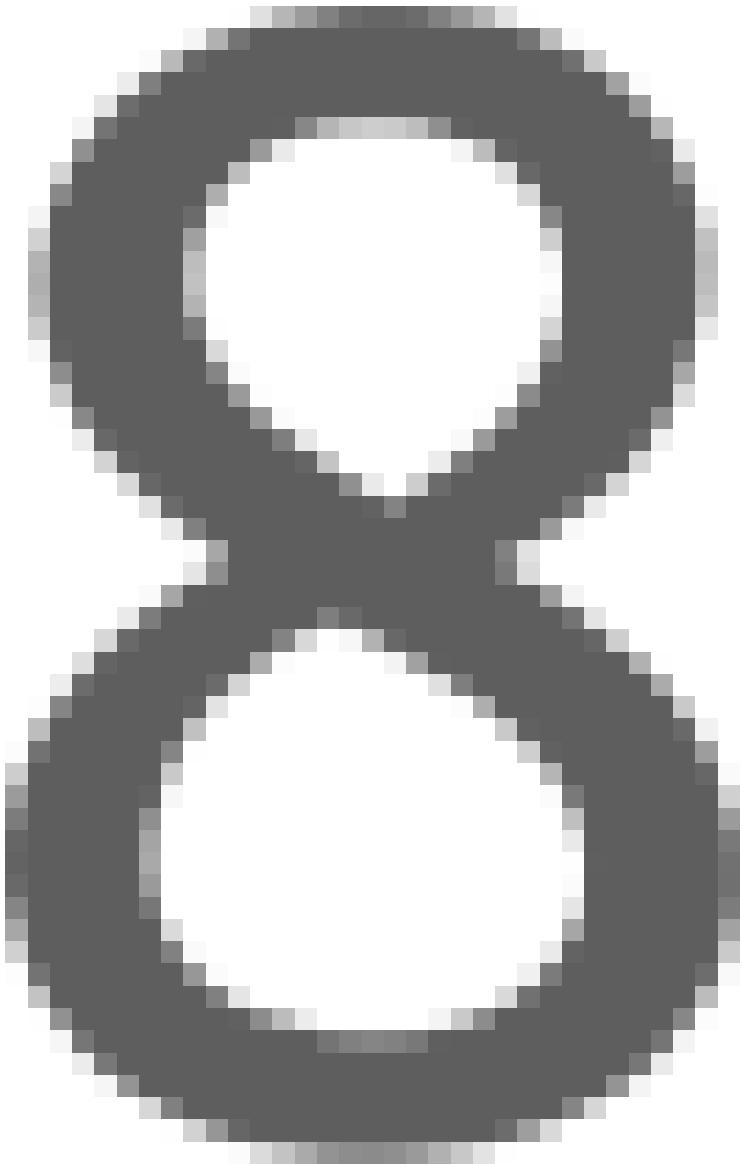




Siddhartha Mishra (MIT/Alfi) Summer School







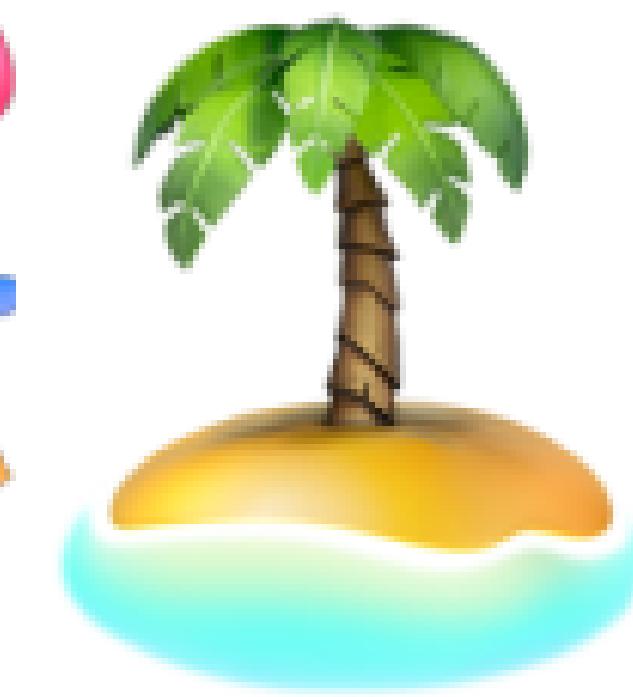
learning the data distribution

How do my generative models learn data?

2. Maximize the likelihood of the model under the training data samples

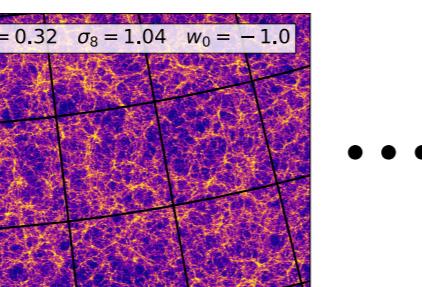
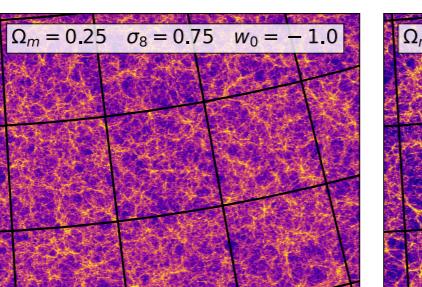
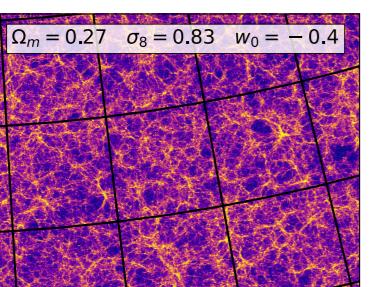
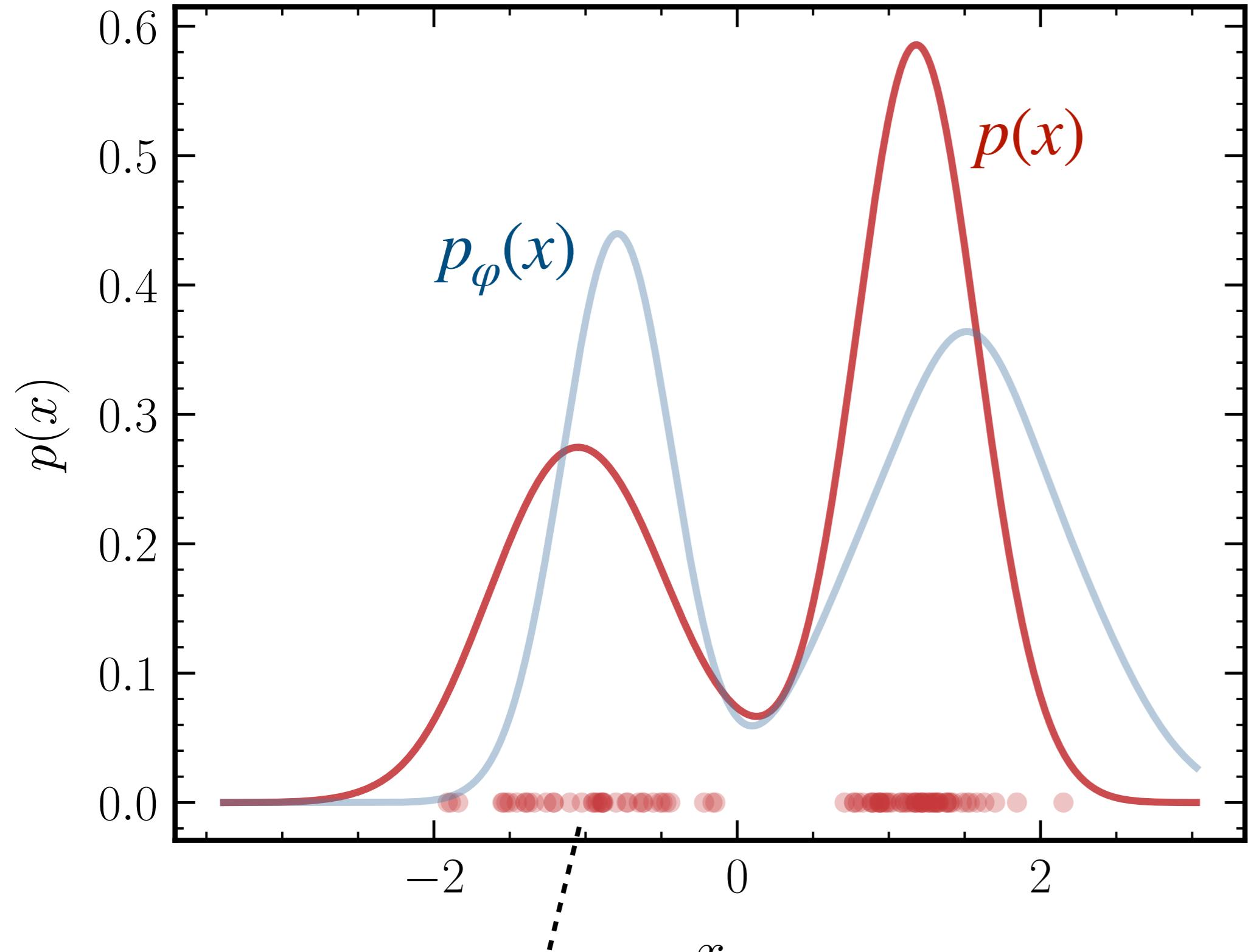
$$\hat{\phi} = \arg \max_{\varphi} \left[ \log p_{\varphi}(\{x\}_{\text{train}}) \right]$$

3. Pô (x)





*Not so fast...*



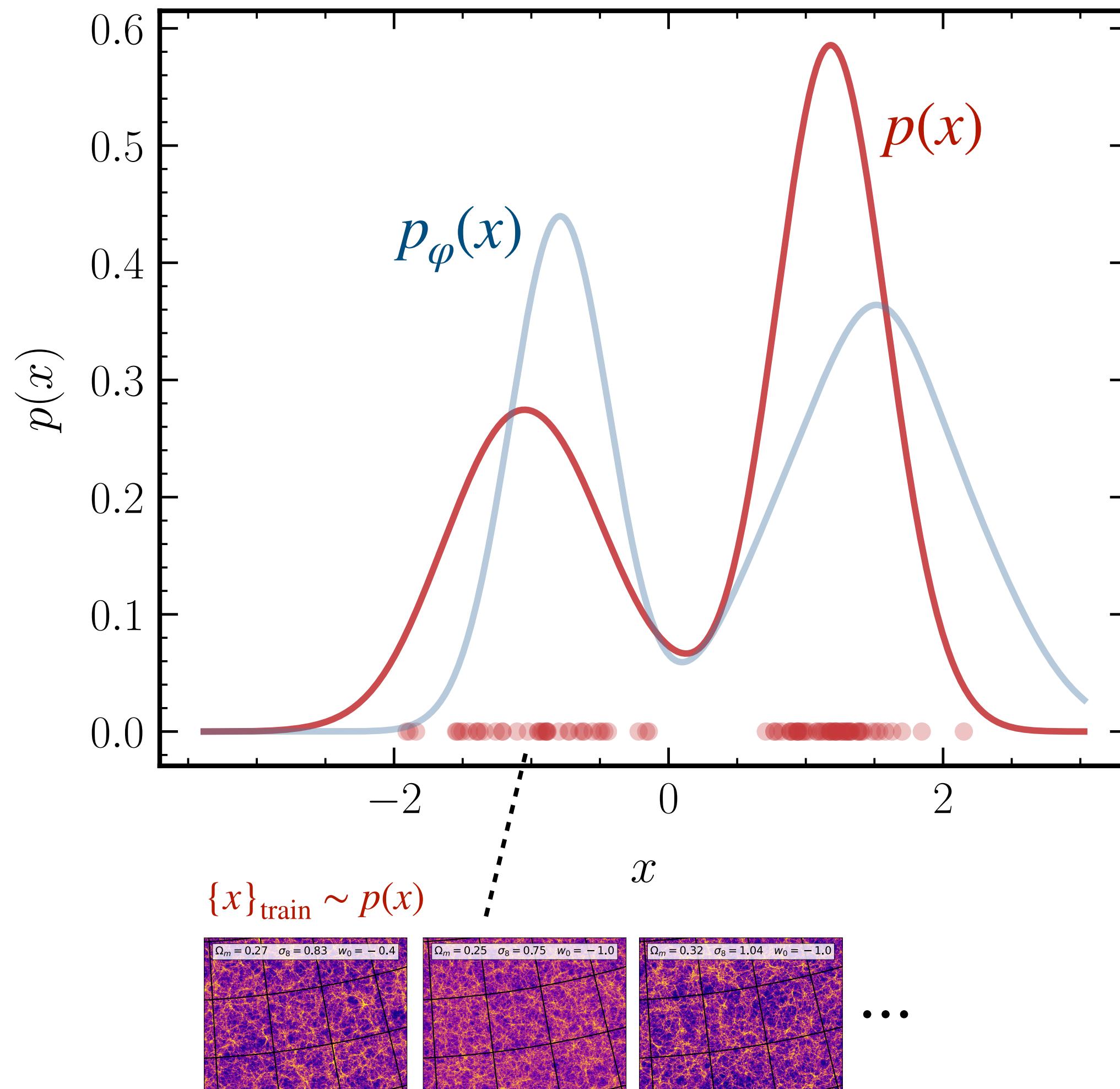
...

## 1. Ingredients:

- A parameterized distribution  $p_\phi(x)$
- Samples from the data distribution  $\{x\}_{\text{train}} \sim p(x)$   
(empirical or simulated)

# Learning the data distribution

I'm sold! *How do I learn a generative model for my data?*



## 1. Ingredients:

- A parameterized distribution  $p_\phi(x)$
- Samples from the data distribution  $\{x\}_{\text{train}} \sim p(x)$   
(empirical or simulated)

## 2. Maximize the likelihood of the model under the training data samples

$$\hat{\phi} = \arg \max_{\varphi} \left[ \log p_\phi \left( \{x\}_{\text{train}} \right) \right]$$

## 3. $p_{\hat{\phi}}(x)$



*Not so fast...*

# The curse of dimensionality

Where is most of the probability mass concentrated in high dimensions?

