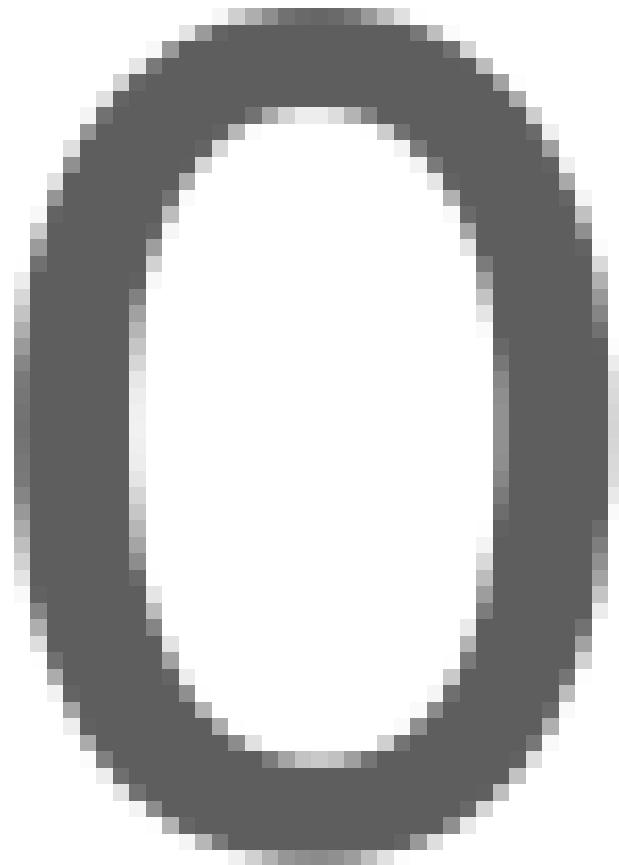
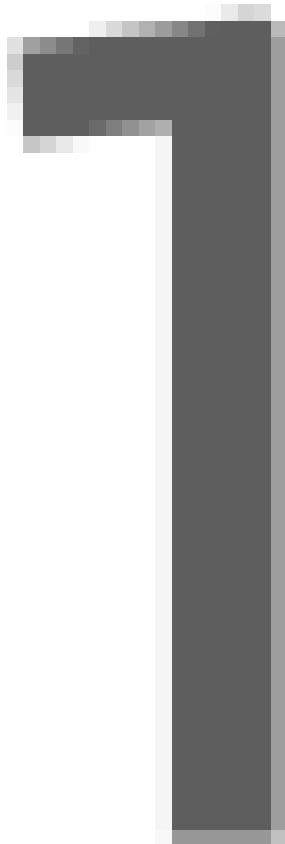


Siddhartha Mishra (MIT/Alfi) Summer School





learning the data distribution

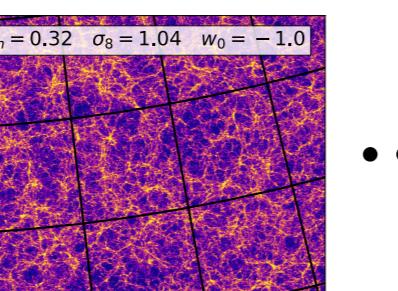
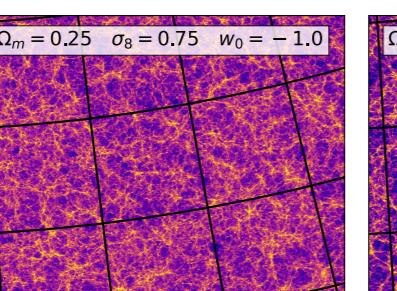
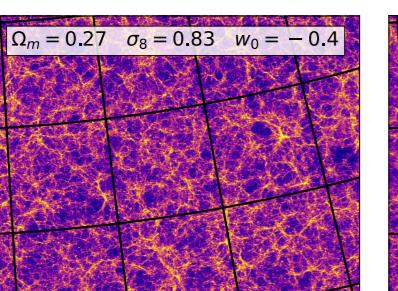
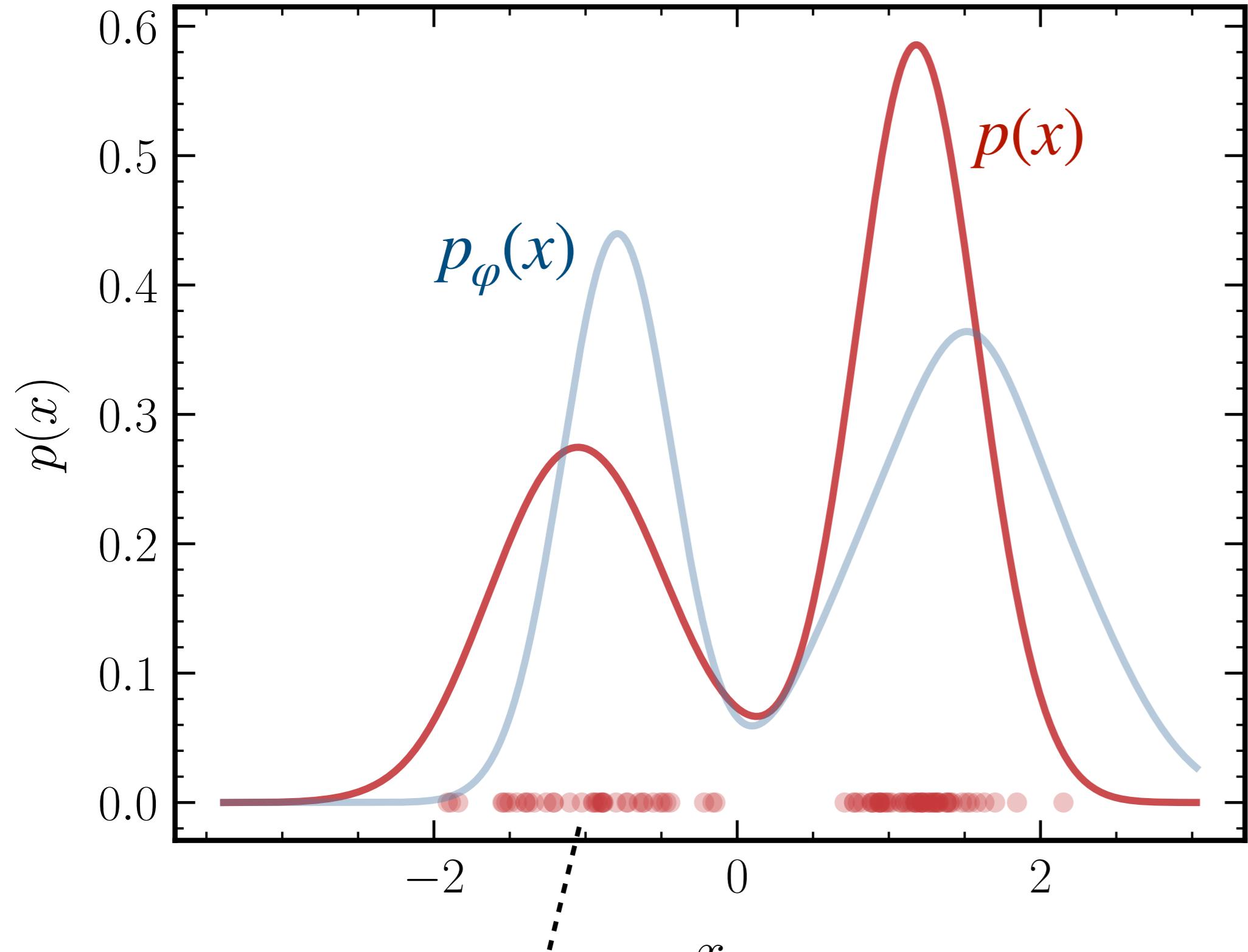
How do my generative models learn data?

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

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



Not so fast...



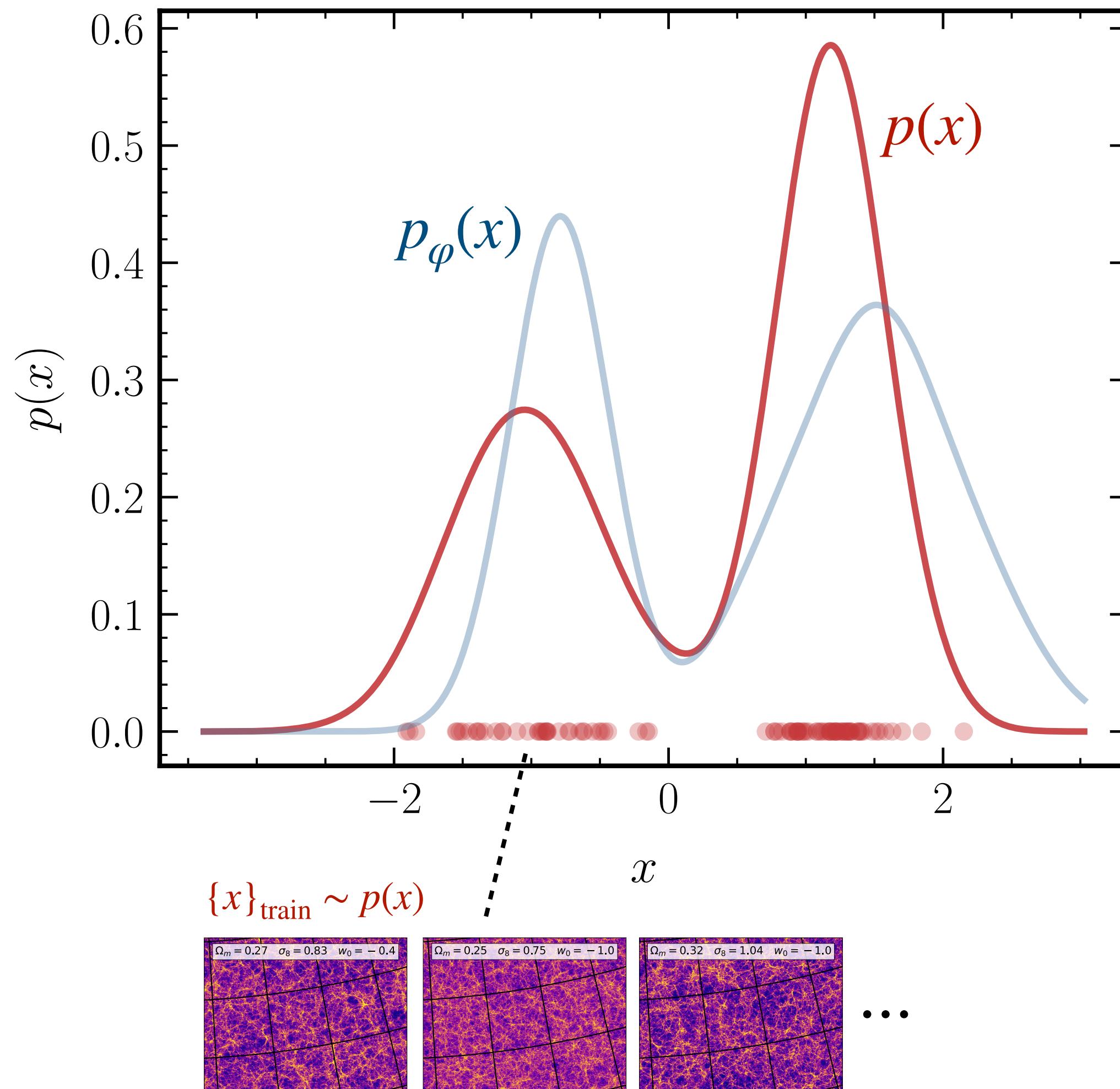
...

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 $p_\varphi(x)$
- 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

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

Not so fast...

The curse of dimensionality

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

