



Siddhant Mishra-Sharma (MIT/AI FI) Summer School



162



Latent-variable modeling

Learn how to find the data distribution

observed variables



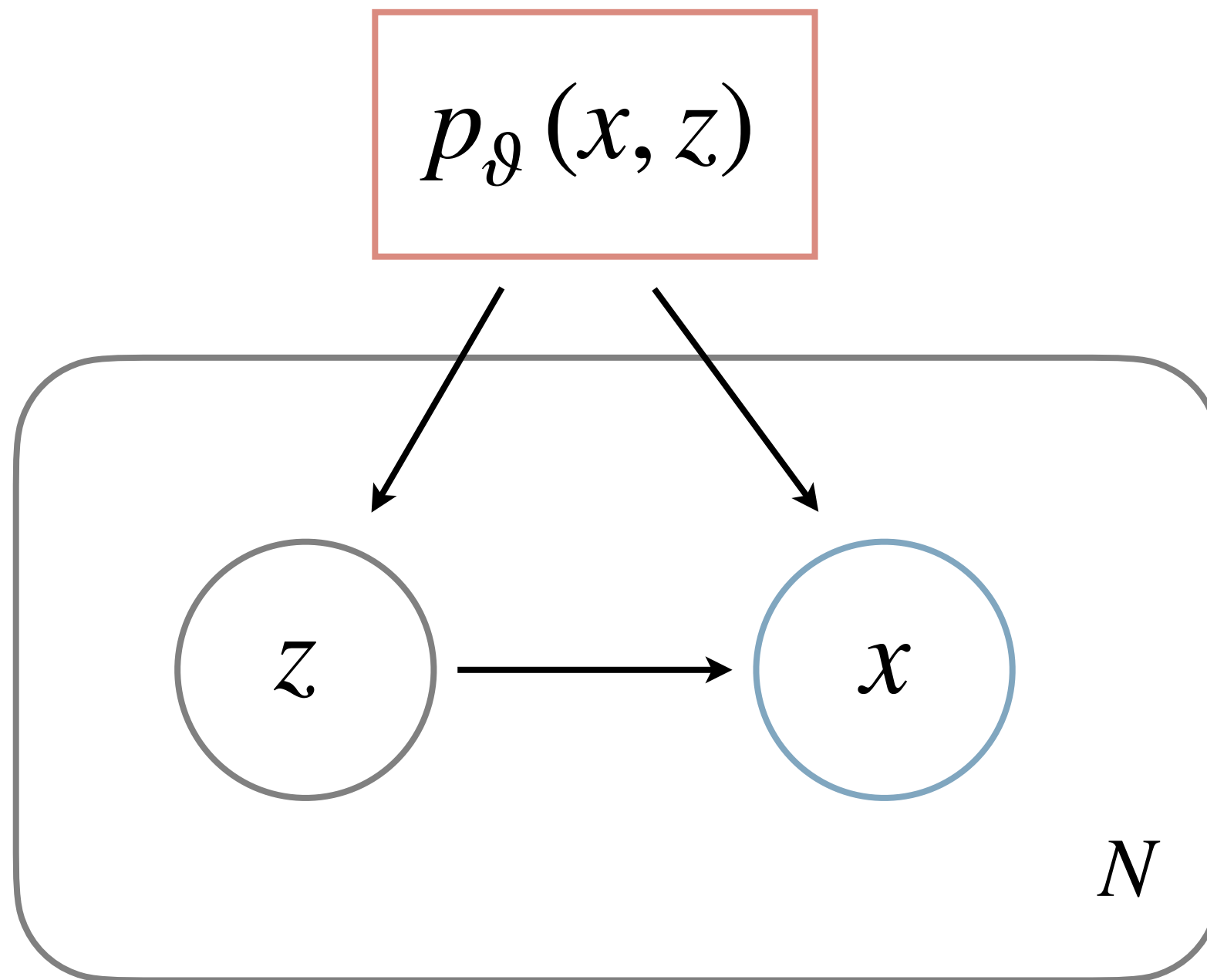
$p_{\vartheta}(x)$



$x$

$N$

Make the problem easier by making it “harder”:  
introduce *joint distribution*  $p_{\theta}(x, z)$



*Latent variables*

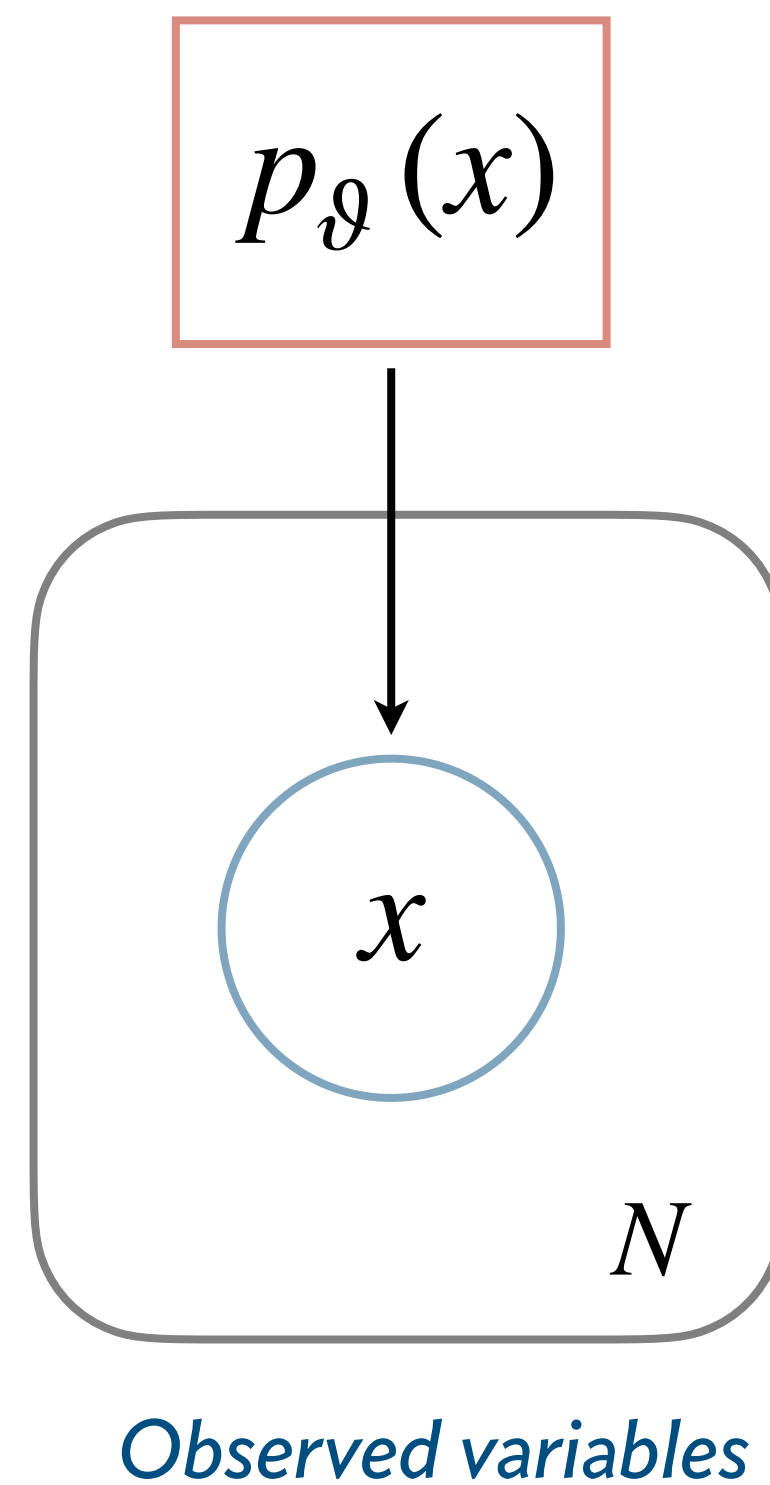
*Observed variables*

# Common factorization:

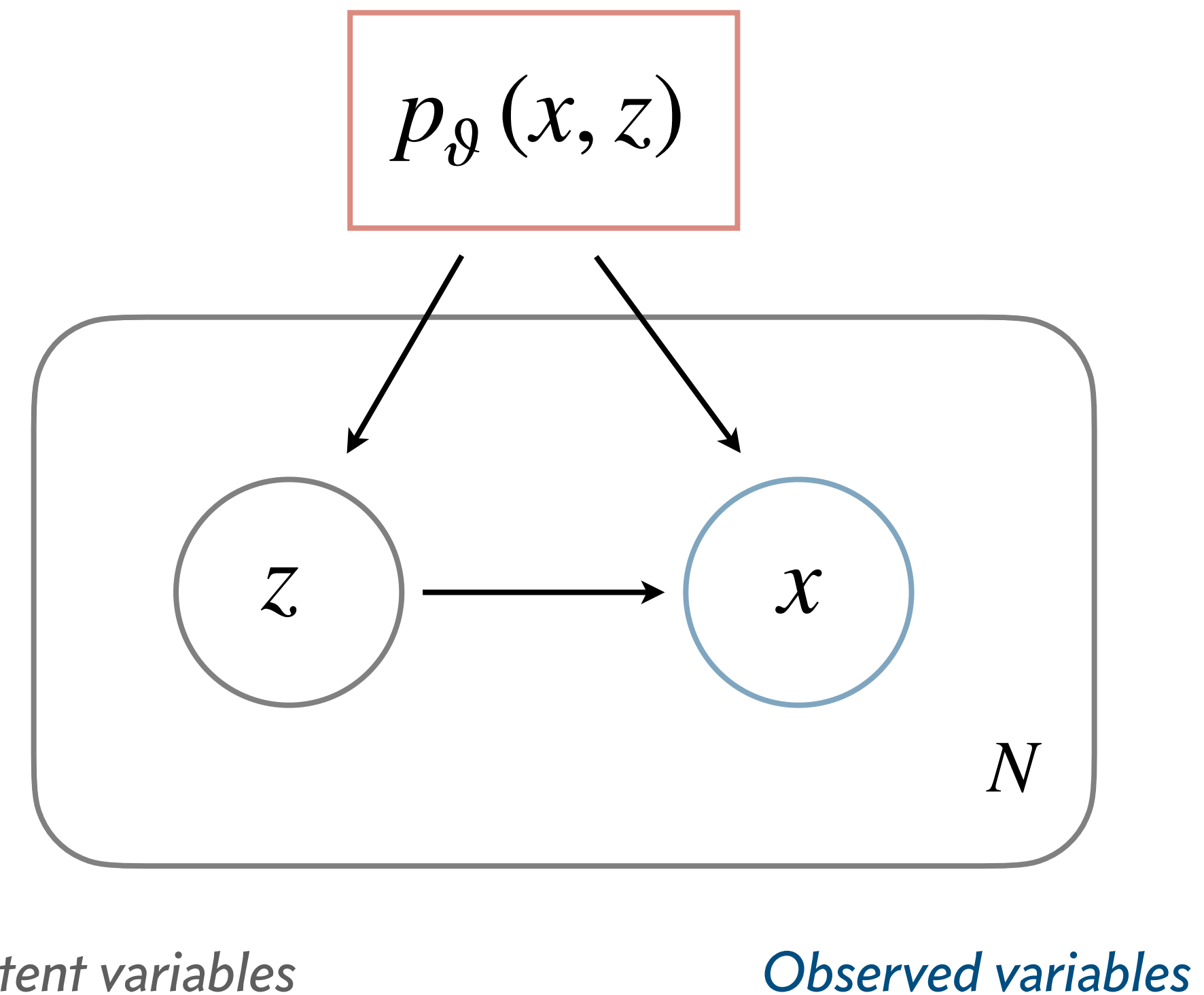
$$p_{\vartheta}(x, z) = p(z) \cdot p_{\vartheta}(x \mid z)$$

# Latent-variable modeling

Learn lower-dimensional structure in the data distribution



Make the problem easier by making it “harder”:  
introduce *joint distribution*  $p_{\theta}(x, z)$



Common factorization:

$$p_{\theta}(x, z) = p(z) \cdot p_{\theta}(x | z)$$

# Latent-variable modeling

## Maximum-likelihood training?

$$\begin{aligned}\vartheta^* &= \arg \max_{\vartheta} p_{\vartheta}(x) \\ &= \arg \max_{\vartheta} \int p_{\vartheta}(x \mid z) p(z) \, dz \\ &= \arg \max_{\vartheta} \left\langle p_{\vartheta}(x \mid z) \right\rangle_{p(z)}\end{aligned}$$

