A Bayesian modeler commits to an a priori *joint distribution*

$$p(\mathbf{y}, \boldsymbol{\theta}) = p(\mathbf{y} \mid \boldsymbol{\theta})p(\boldsymbol{\theta}) = p(\boldsymbol{\theta} \mid \mathbf{y})p(\mathbf{y})$$

Likelihood x Prior

Posterior x Marginal Likelihood

Data *(observed)*

Parameters *(unobserved)*

What is the problem with “vague” priors?