

Scientific questions

Experimental design,
Database queries,
Bioinformatic pipelines

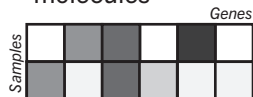
Observed data



- Amino acid sequence
- DNA sequence

N1CCN(C.....)

- SMILES encoded molecules



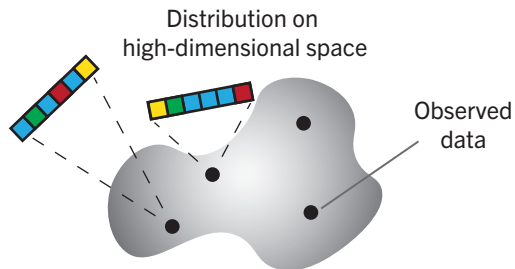
- Gene expression

⋮

- Microscopy images
- Chromatin data (Hi-C)

Box's loop

Deep generative model



Model

Probabilistic models,
Generative networks,
Domain-specific components

Update model

Inference

Evidence lower bound (VAE),
adversarial inference (GAN),
with automatic differentiation libraries
to learn neural network parameters

Apply model

Criticism

Task-specific benchmarks,
posterior predictive checks,
held-out metrics

