Does Model Match the Data Generating Process?

- Every method involves assumptions about the data distribution

i.e. the data generating process (DGP)

- **Likelihood-based methods** are based on a probabilistic model for the data

- assumptions are explicit $\Rightarrow$ tend to be more interpretable

- **Algorithmic methods** specify an algorithm or objective function to optimize

- assumptions are implicit $\Rightarrow$ tend to be less interpretable

- Even the simplest method performs optimally if its assumptions match the DGP

- But if little is known about the DGP, then a more flexible method may be preferable

- Note: some methods are kind of in-between

Does Model Match the Data Generating Process?

- Every method involves assumptions about the data distribution
  i.e. the data generating process (DGP)

- **Likelihood-based methods** are based on a probabilistic model for the data
  - assumptions are explicit $\Rightarrow$ tend to be more interpretable

- **Algorithmic methods** specify an algorithm or objective function to optimize
  - assumptions are implicit $\Rightarrow$ tend to be less interpretable

- Even the simplest method performs optimally if its assumptions match the DGP

- But if little is known about the DGP, then a more flexible method may be preferable

- Note: some methods are kind of in-between

Likelihood-based vs. Algorithmic method?