1. The Garden of Forking Data

* Workflow and Questions
  + How should we use the sample?
  + How to produce a summary?
  + How to represent uncertainty?
  + Workflow
    - Define generative model of the sample
    - Define a specific estimand
    - Design a statistical way to produce estimate
    - Test 3 using 1
    - Analyze sample, summarize
* Bayesian data analysis
  + For each possible explanation of the sample, count all the ways the sample could happen. Explanations with more ways to produce the sample are more plausible.
* Test before you estimate
  + Code a generative simulation
  + Code an estimator
  + Test 2 with 1
* Simulating
  + Test the estimator where the answer is known
  + Explore different sampling designs
  + Develop intuition for sampling and estimation
* The Beta Distribution (51:00)
* Rules of Bayesian updating
  + No minimum sample size
  + Shape of the posterior distribution embodies sample size
  + There are no point estimates. The estimate is the posterior distribution.
  + There is no one true interval. Nothing special happens at the end of an interval because the interval is arbitrary.
* From posterior to prediction
  + Implications of model depend upon entire posterior. Must average any inference over entire posterior. This usually requires integral calculus or we can just take samples from the posterior.
* That an arbitrary interval contains an arbitrary value is not meaningful. Use the whole distribution.

Uncertainty 🡪 Causal Model 🡪 Implications

* Posterior Predictive Distribution: Based on what we know so far, what do we expect to happen if we continue to take samples from the distribution?
* Bayesian Modesty: No guarantees except logical. Any framework selling you more is hiding assumptions.