Reproducible workflow

1. Document everything
2. Do (almost) everything using code
3. Use open source software & file formats
4. Organize files in one location
5. Track changes to files
6. Archive final working versions
7. Backup all files

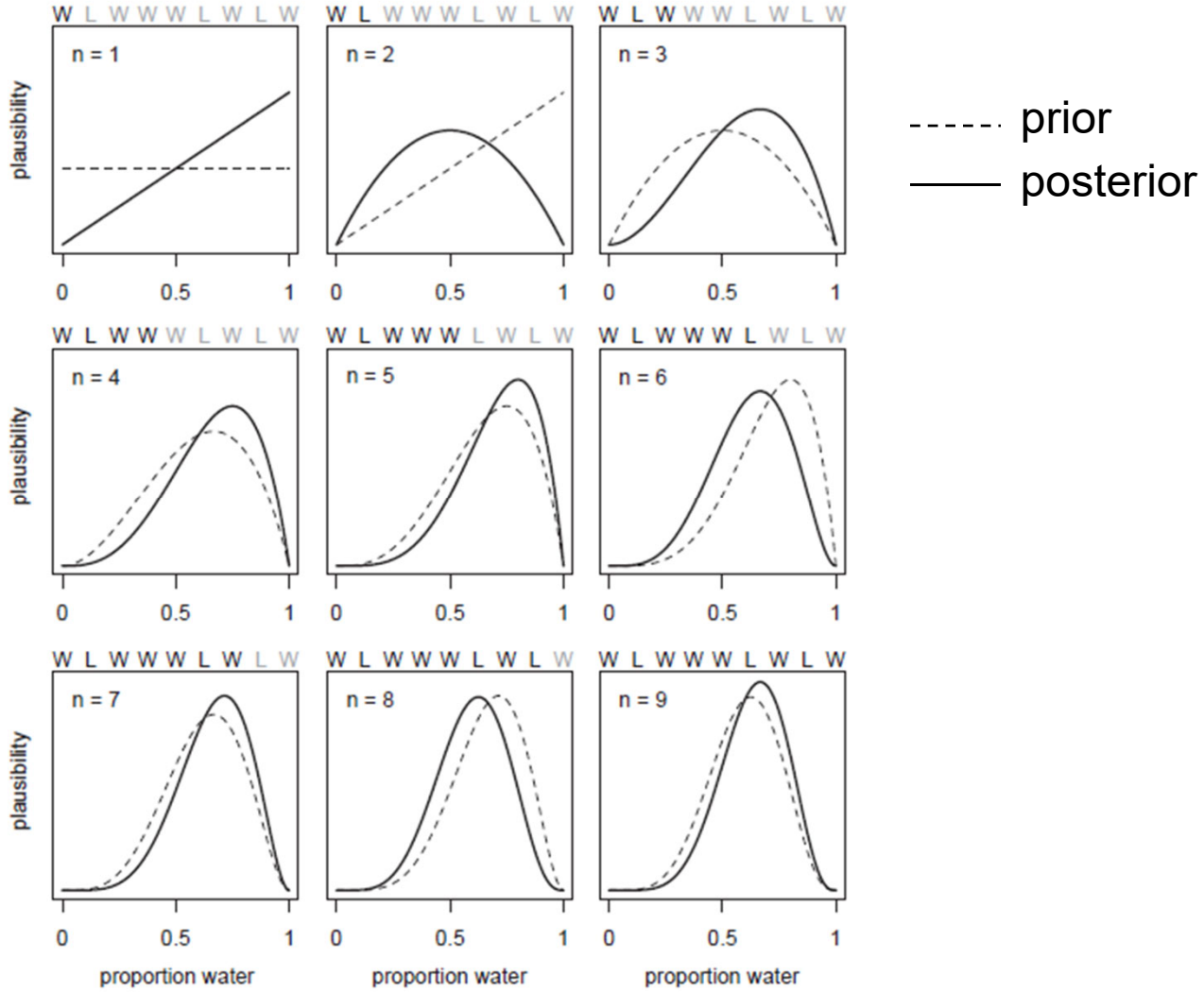
Markdown editors

- Marktext – my current editor
- Rstudio
- VSCode - Microsoft
- VSCodium – version with no telemetry

Main concepts McElreath 2

- Golem = algorithm
- Likelihood: counting all the ways data could have happened
- Bayesian updating: prior x likelihood
  - using counts
  - using probabilities
  - using distributions

Bayesian updating



Components of model

1) Likelihood
   - "data story" = data generating process
   - from first principles, or “off the shelf”
2) Parameters
   - quantities that don't change
   - to be estimated
3) Prior distribution
4) Posterior distribution (inference)
   - histogram *is* the posterior

Bayesian inference

$$P(B|A) = \frac{P(B)P(A|B)}{P(A)}$$

Bayes' rule for two events A, B

"Posterior"

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

Model

Data

"Prior"

Likelihood

Total probability of the data

Apply Bayes' rule to convert the likelihood into what we really want to know: the probability of the model given the data

P(y): probability added up or integrated over all of the models

$$P(y) = \sum_{\theta} P(\theta)P(y|\theta)$$

Discrete parameter

$$P(y) = \int P(\theta)P(y|\theta)\, d\theta$$

Continuous parameter

Grid approximation

Algorithm

load data

define grid of parameter values

for each parameter value

    compute prior probability

    compute likelihood

    numerator = prior x likelihood

denominator = sum of numerators

for each parameter value

    posterior probability = numerator / denominator

plot posterior probability vs parameter values

unstandardized posterior

total probability

posterior distribution