Today

- Recap & questions from homework
- Pair programming

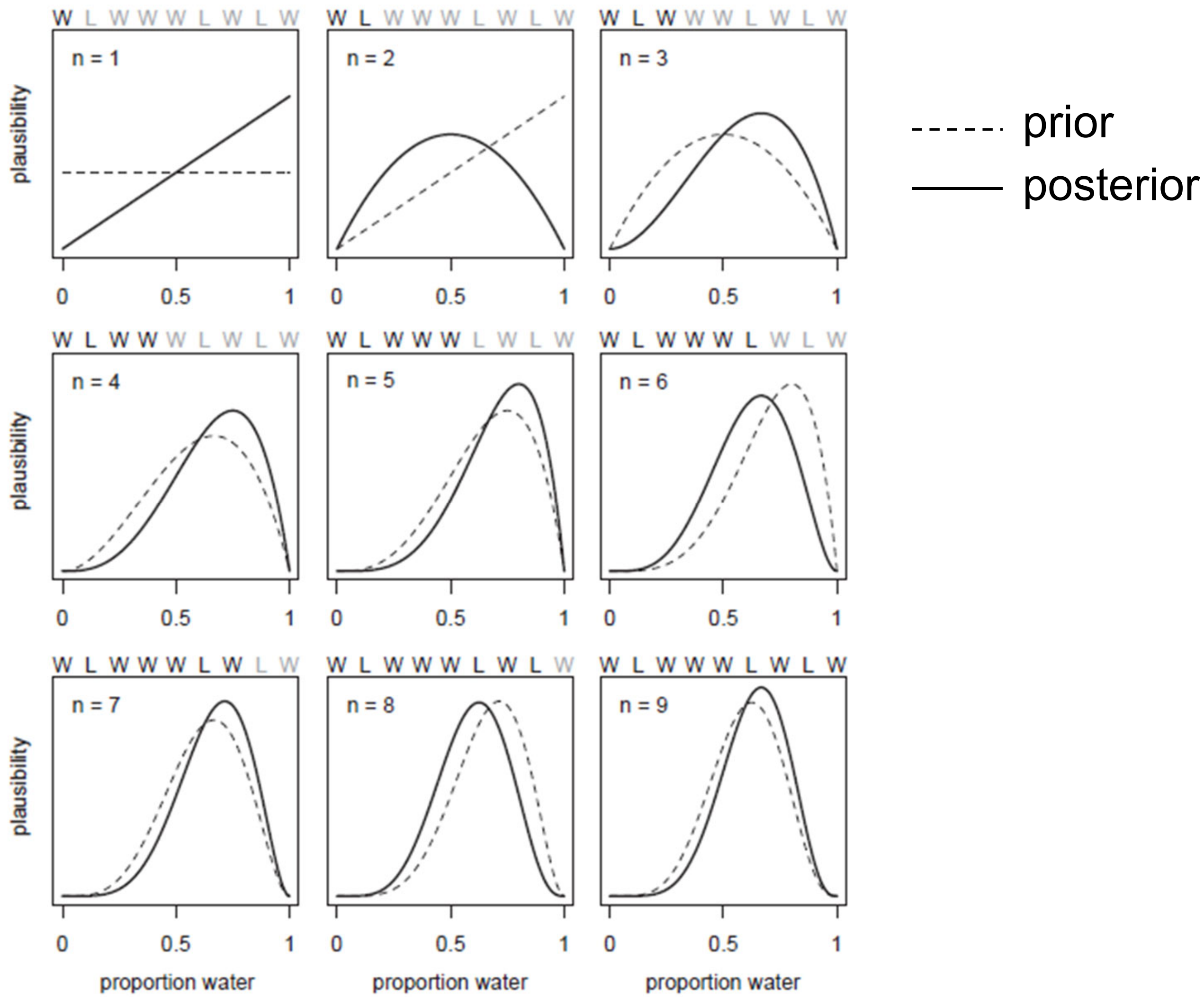
GitGUI & Gitk

- MacOS
  - install via homebrew
  - https://git-scm.com/downloads/mac
  - `$ brew install git-gui`

Main concepts McElreath 2

- Golem = algorithm
- Small world (model) vs large world (reality)
- Likelihood: counting all the ways data could have happened
- Bayesian updating: prior x likelihood
  – using counts (marbles)
  – using probabilities (marbles; divide by sum of numerator)
  – using distributions (p_water via Earth toss)

Bayesian updating



Pair programming

- Chapter 2
  - Coding Bayesian updating
  - Questions: 2M1, 2M2

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

Bayesian inference

Unstandardized posterior

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

Total probability standardizes the posterior to a probability (discrete parameter) or probability density (continuous parameter)

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

Discrete parameter

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

Continuous parameter

Posterior distribution algorithm

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

Algorithm
load data
for each parameter value
    unstandardized posterior = prior * likelihood
calculate the total probability
for each parameter value
    posterior probability =
        unstandardized posterior / total probability
plot posterior probability vs parameter values

posterior distribution

Discrete parameter

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

Algorithm
load data
for each parameter value
    unstandardized posterior = prior * likelihood
total probability = sum(unstandardized posteriors)
for each parameter value
    posterior probability =
        unstandardized posterior / total probability
plot posterior probability vs parameter values

posterior distribution

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

probability mass function (pmf)

Discrete parameter

- Uncommon
- Usually a latent (hidden) state
- Examples
  - male/female in population model
  - occupied/unoccupied in species occupancy model

Continuous parameter

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

Exact algorithm
load data
define grid of parameter values
for each parameter value
    unstandardized posterior = prior * likelihood
total probability = integral(unstandardized posterior function)
for each parameter value
    posterior Pr density =
        unstandardized posterior / total probability
plot posterior Pr density vs parameter values

area under curve

posterior distribution

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

probability density function (pdf)

Continuous parameter

Grid approximation algorithm

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

load data
define grid of parameter values with resolution r
for each parameter value
    unstandardized posterior = prior * likelihood
total probability = sum(unstandardized posteriors) * r
for each parameter value
    posterior Pr density =
        unstandardized posterior / total probability
plot posterior Pr density vs parameter values

approximate area under curve

posterior distribution

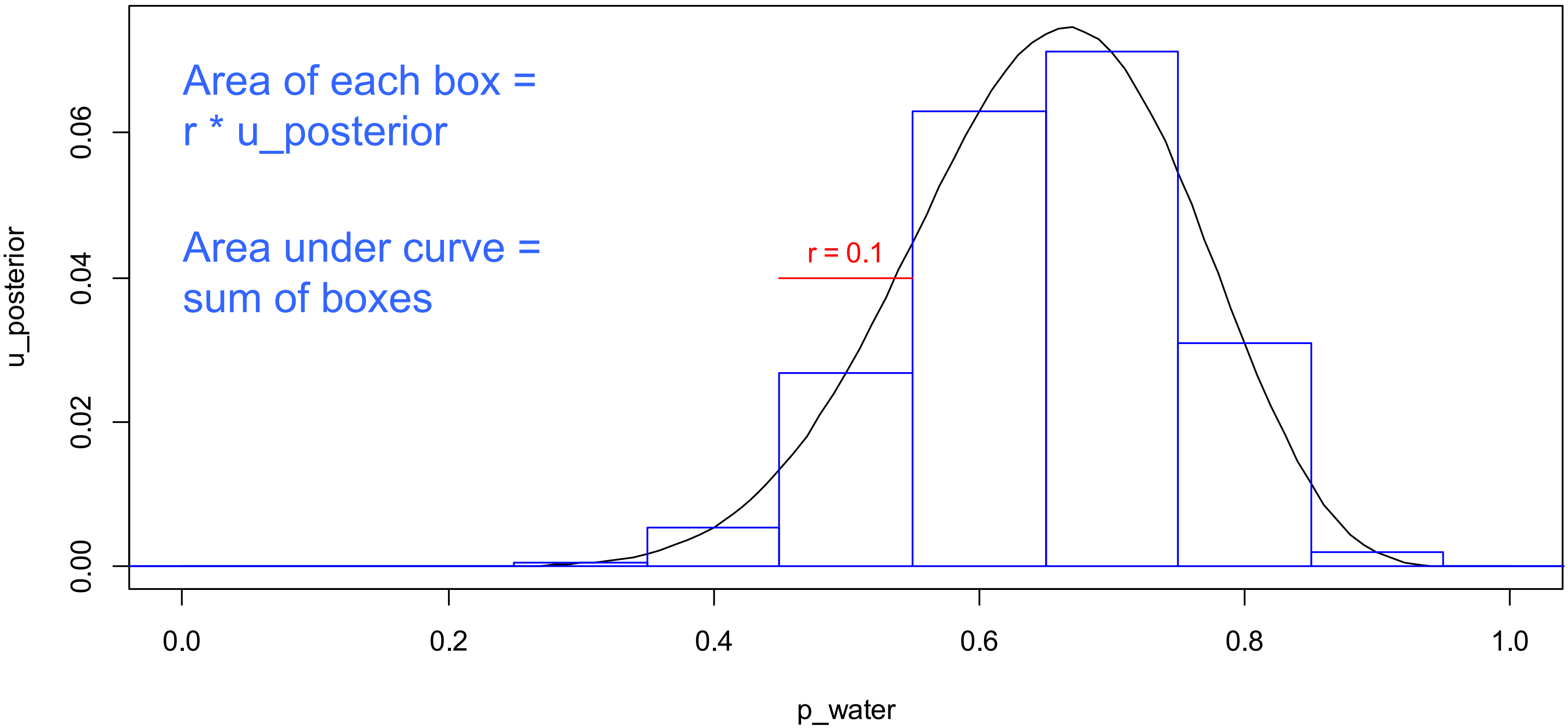
numerical integration

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

probability density function (pdf)

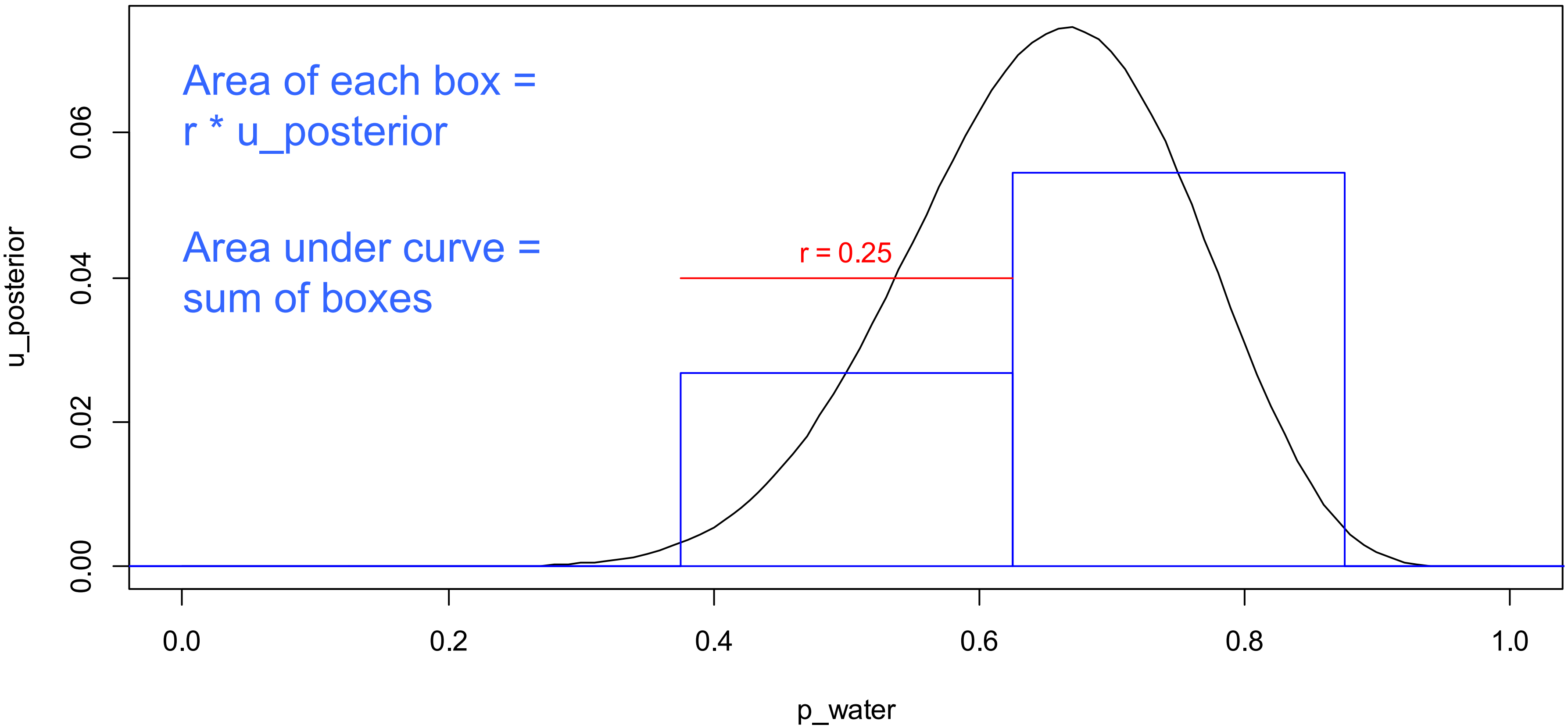
Numerical integration

Approximating area under the curve



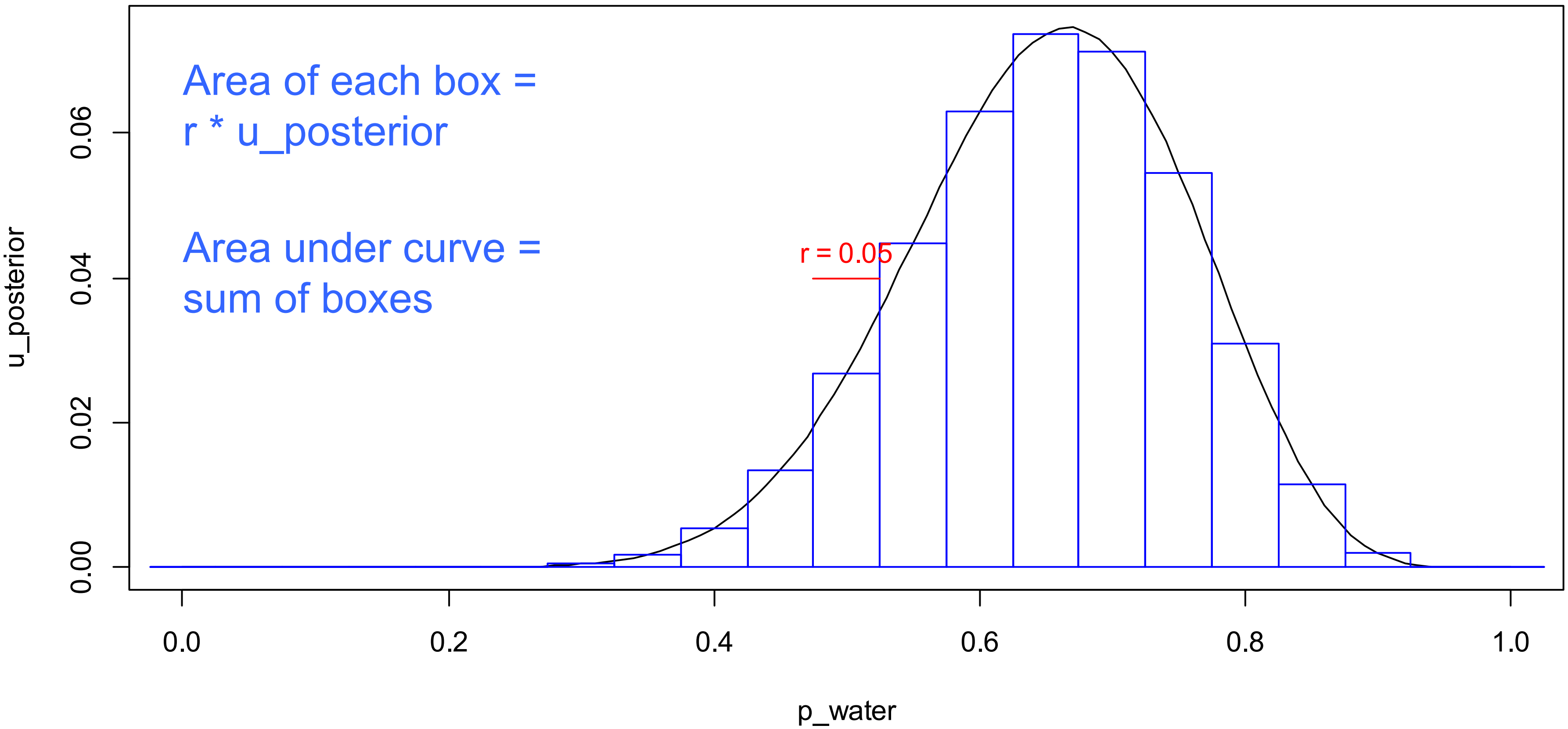
Numerical integration

Approximating area under the curve



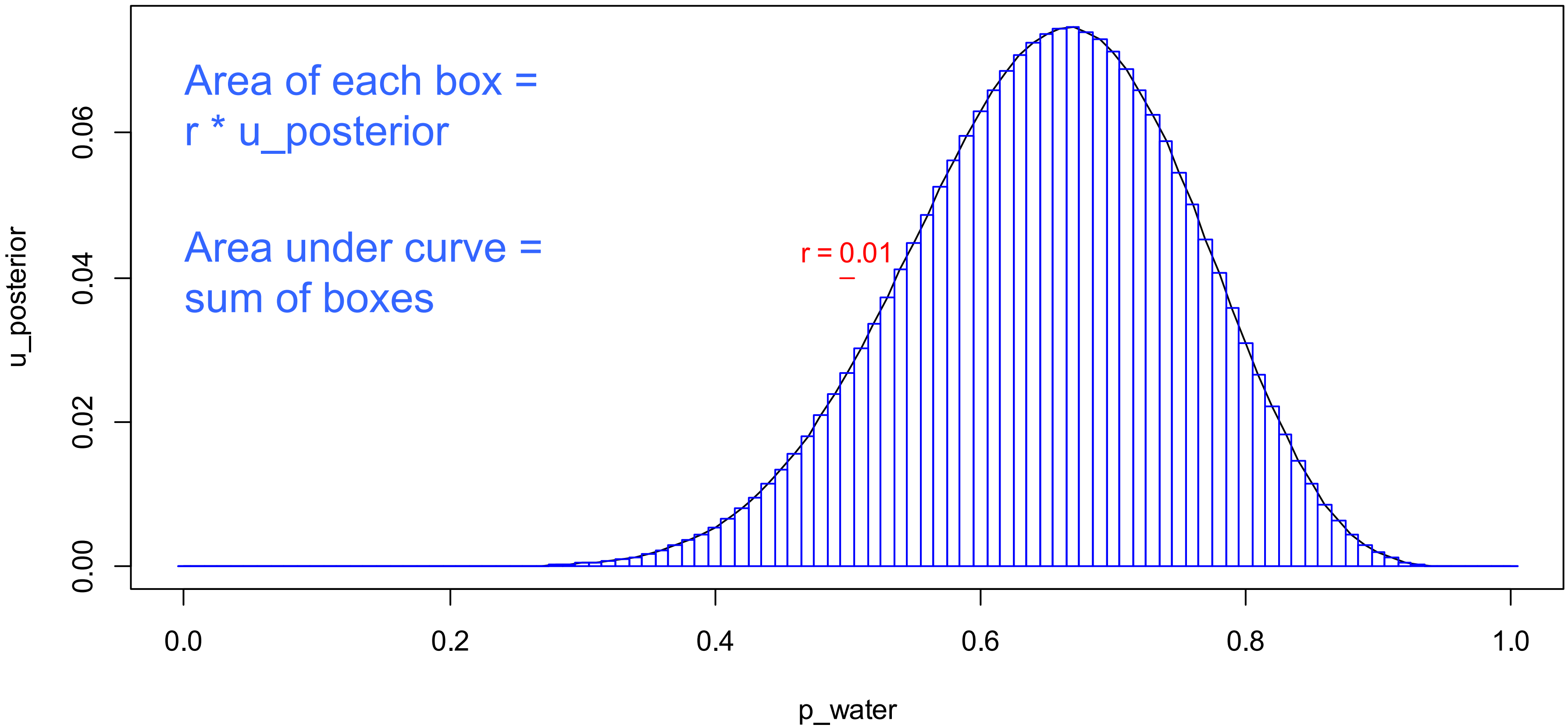
Numerical integration

Approximating area under the curve



Numerical integration

Approximating area under the curve



Numerical integration

Approximating area under the curve