# Today

- Bayesian inference

# Bayesian inference

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

Bayes' rule for two events A, B

"Posterior"

$$\underbrace{P(\theta|y)}_{\text{Model, Data}} = \frac{\overbrace{P(\theta)}^{\text{"Prior"}}\overbrace{P(y|\theta)}^{\text{Likelihood}}}{\underbrace{P(y)}_{\text{Total probability of the data}}}$$

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

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

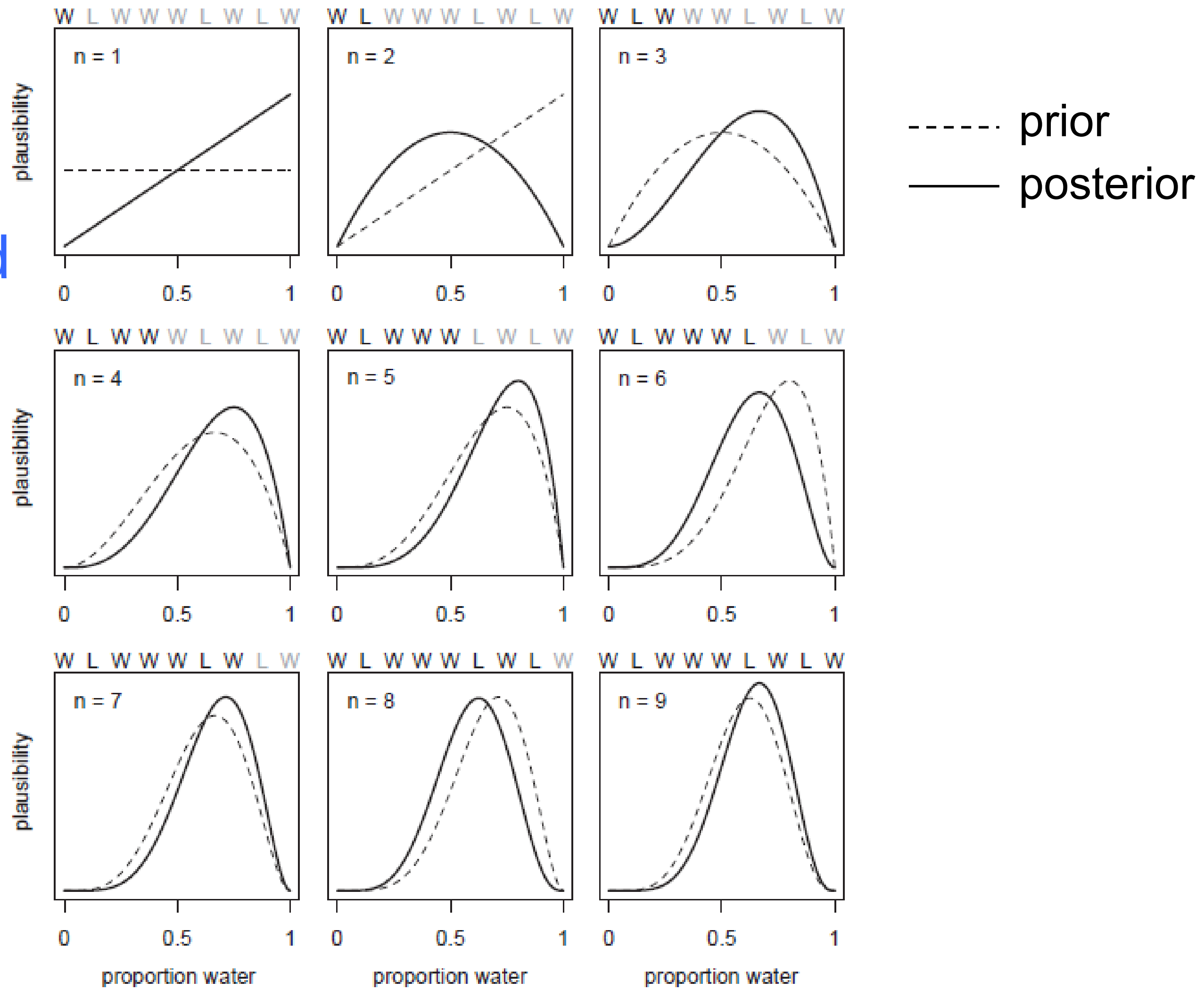
Discrete parameter

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

Continuous parameter

# Bayesian updating

Prior x likelihood



# Components for inference

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 for each parameter
4) Posterior distribution (inference)
   - histogram *is* the posterior

# 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

$\boxed{\text{total probability = integral(unstandardized posterior function)}}$

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

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

Grid approximation algorithm

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

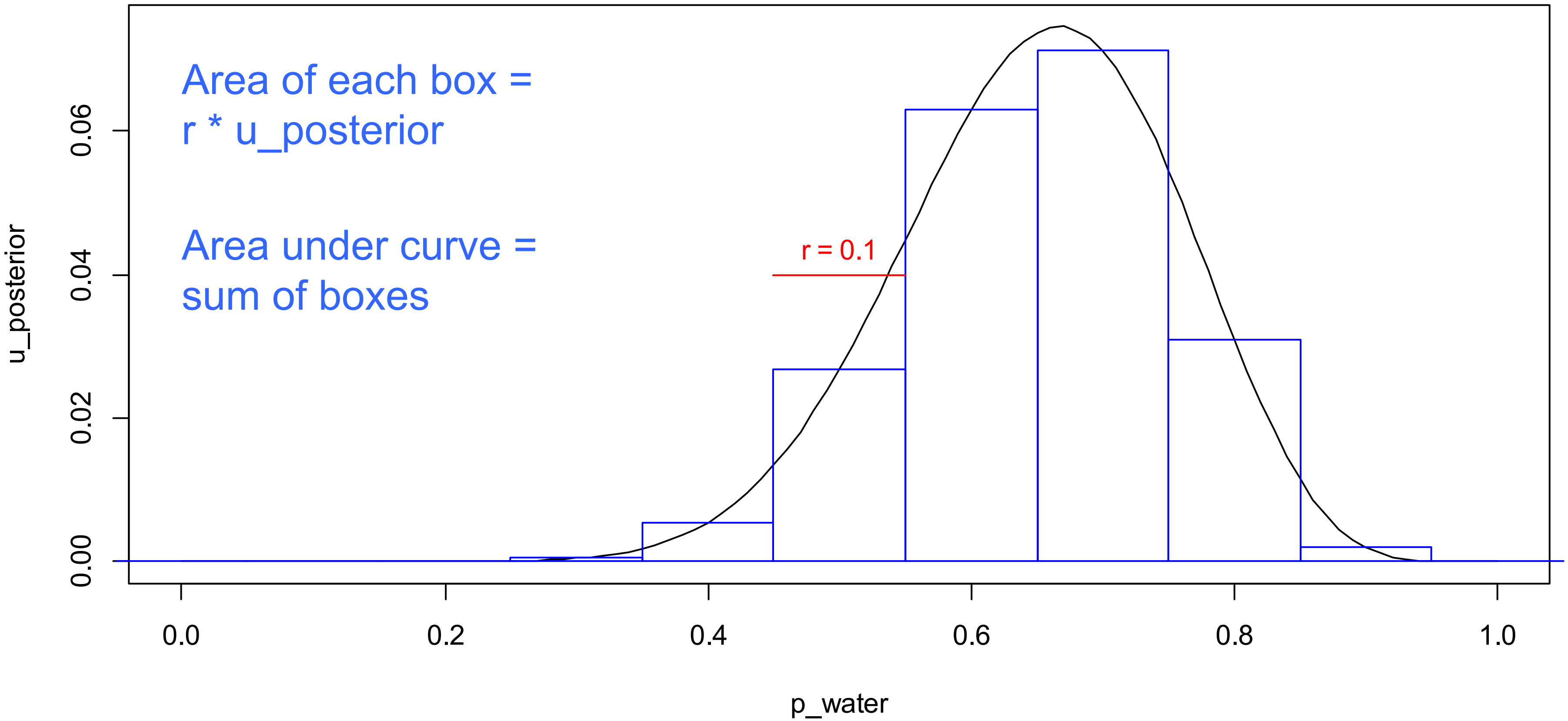
numerical integration

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

probability density function (pdf)

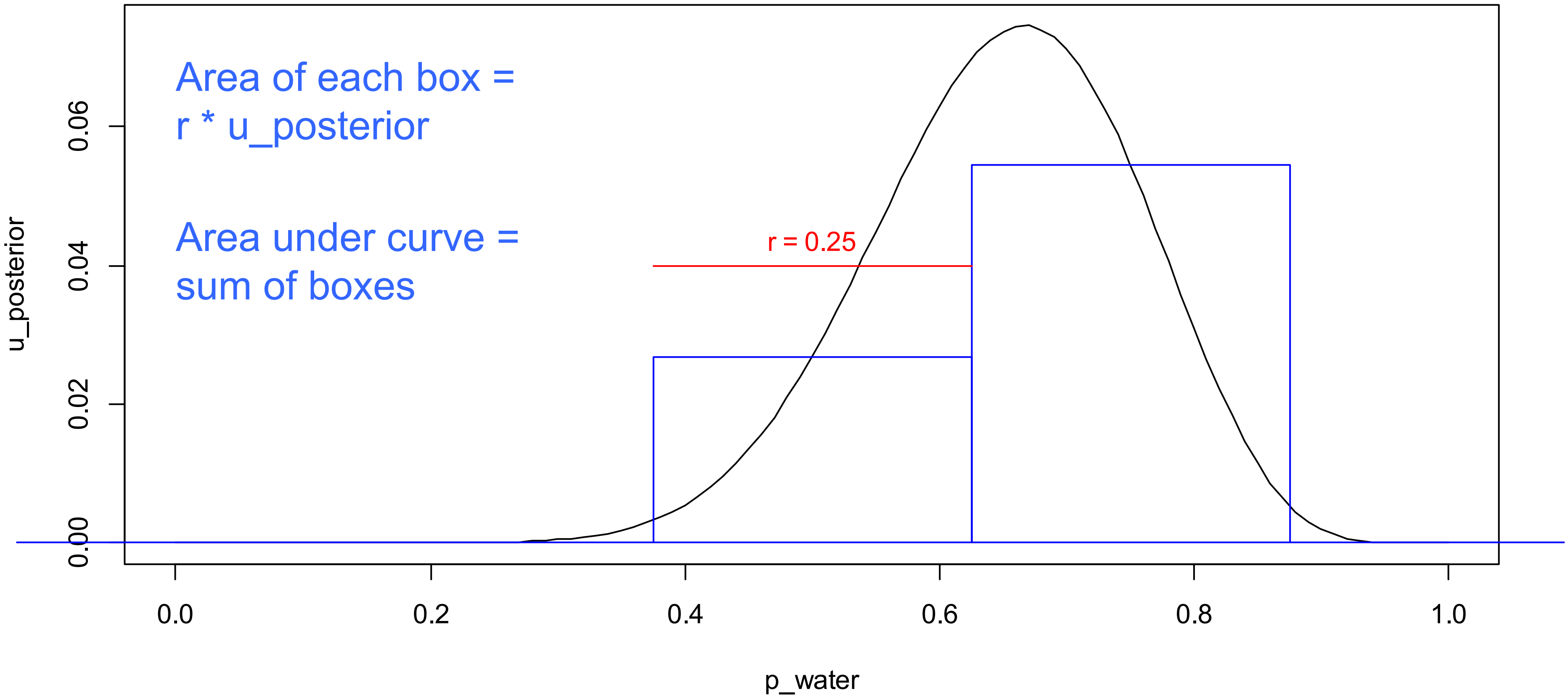
# Numerical integration

Approximating area under the curve



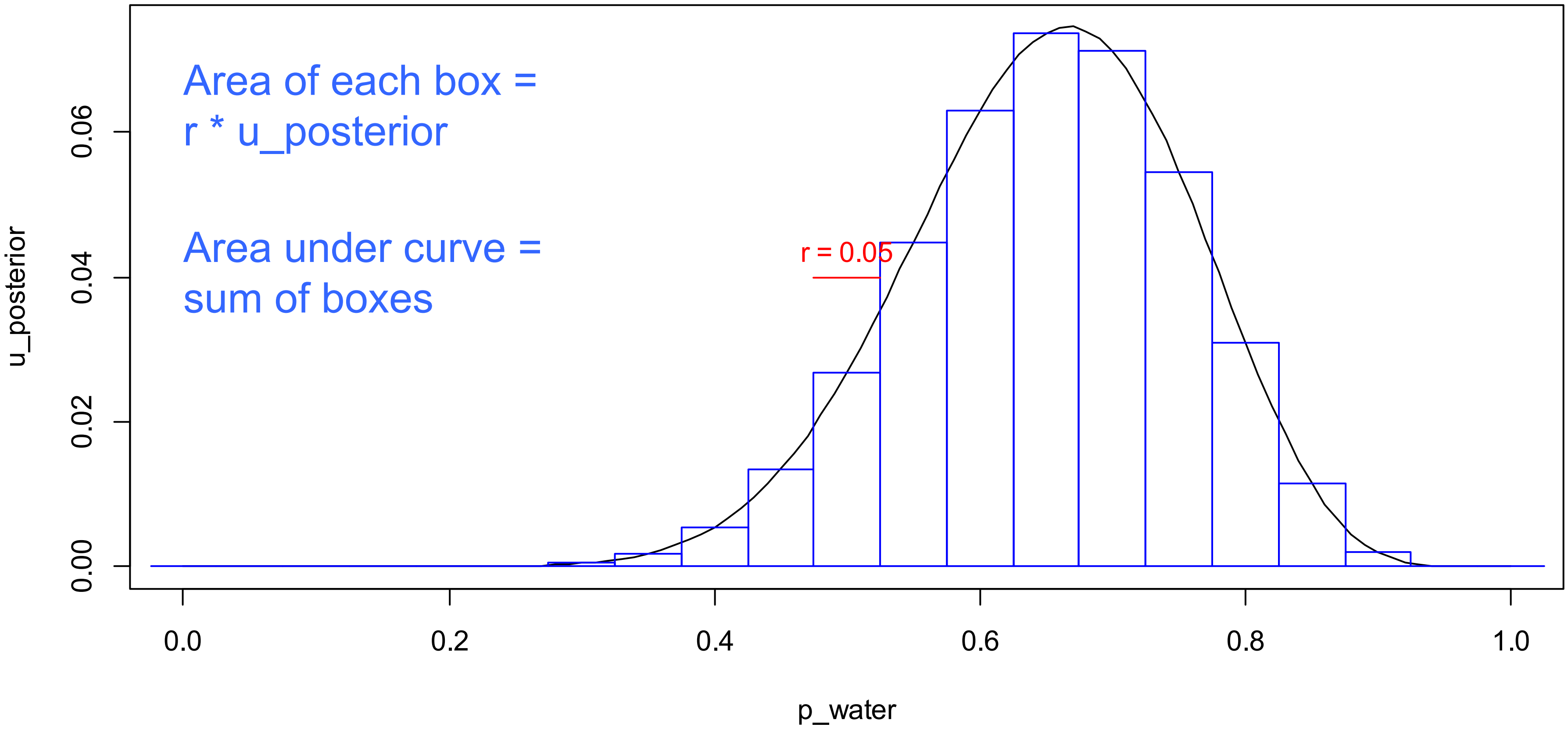
# Numerical integration

Approximating area under the curve



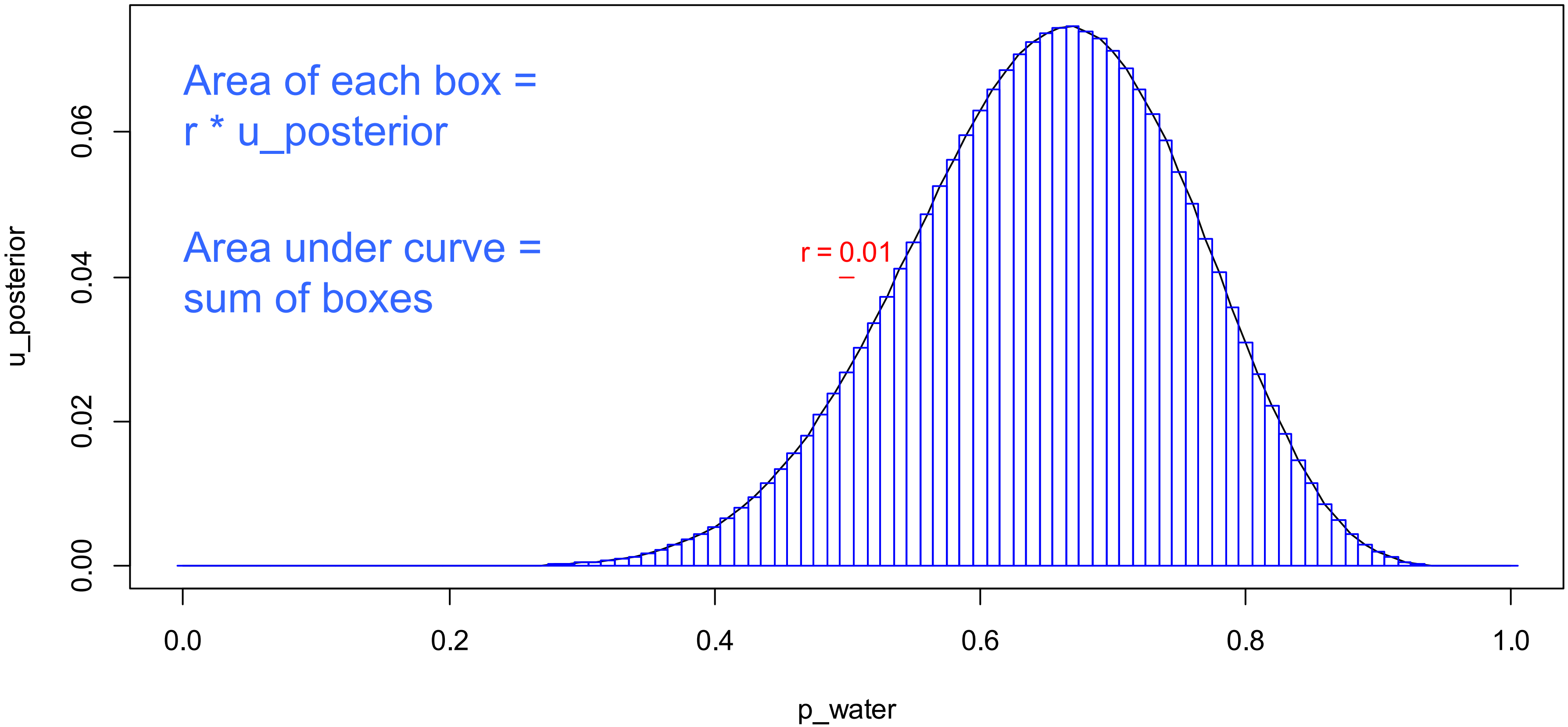
# Numerical integration

Approximating area under the curve



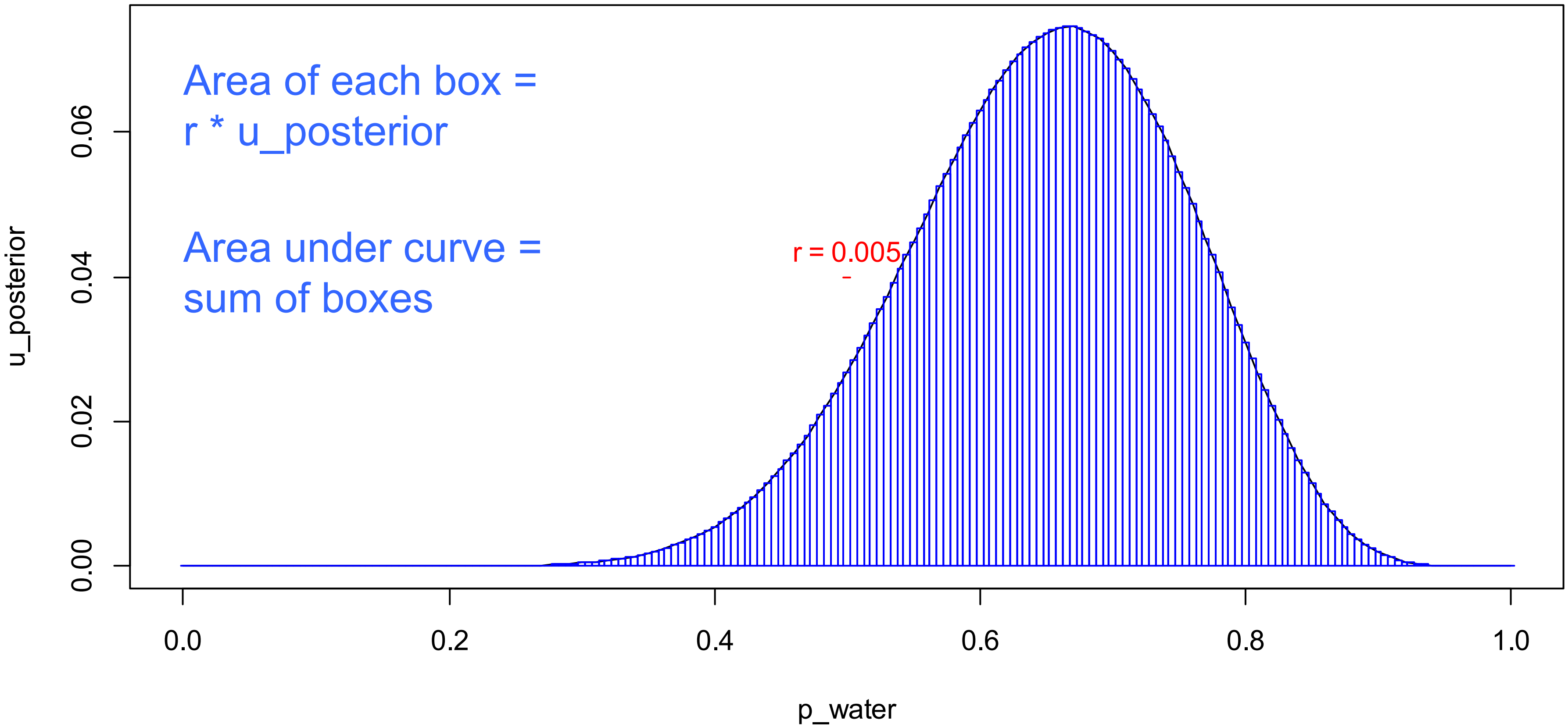
# Numerical integration

Approximating area under the curve



# Numerical integration

Approximating area under the curve



# Main messages from McElreath Ch 4

- Language for describing models

function list model definition

e.g. a linear model

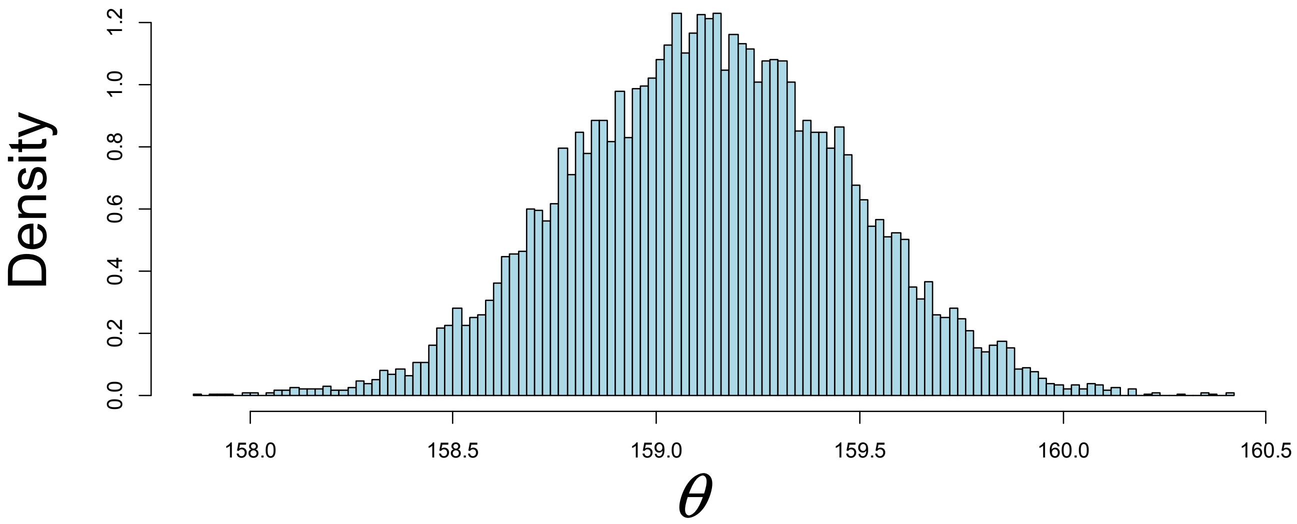
$$y_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \beta_0 + \beta_1 x_i$$
$$\beta_0 \sim \text{Normal}(178, 100)$$
$$\beta_1 \sim \text{Normal}(0, 10)$$
$$\sigma \sim \text{Uniform}(0, 50)$$

```
flist <- alist(
      y ~ dnorm( mu , sigma ),
      mu <- beta_0 + beta_1 * x,
      beta_0 ~ dnorm(178, 100),
      beta_1 ~ dnorm(0, 10),
      sigma ~ dunif(0, 50)
)
sampost(flist, data=df)
```

Sample from the posterior distribution of the parameters

# Main messages from McElreath Ch 4

- Histogram is the posterior distribution




- Obtain all our inferences (means, credible intervals, prediction intervals) from the posterior samples of parameters

# Posterior samples

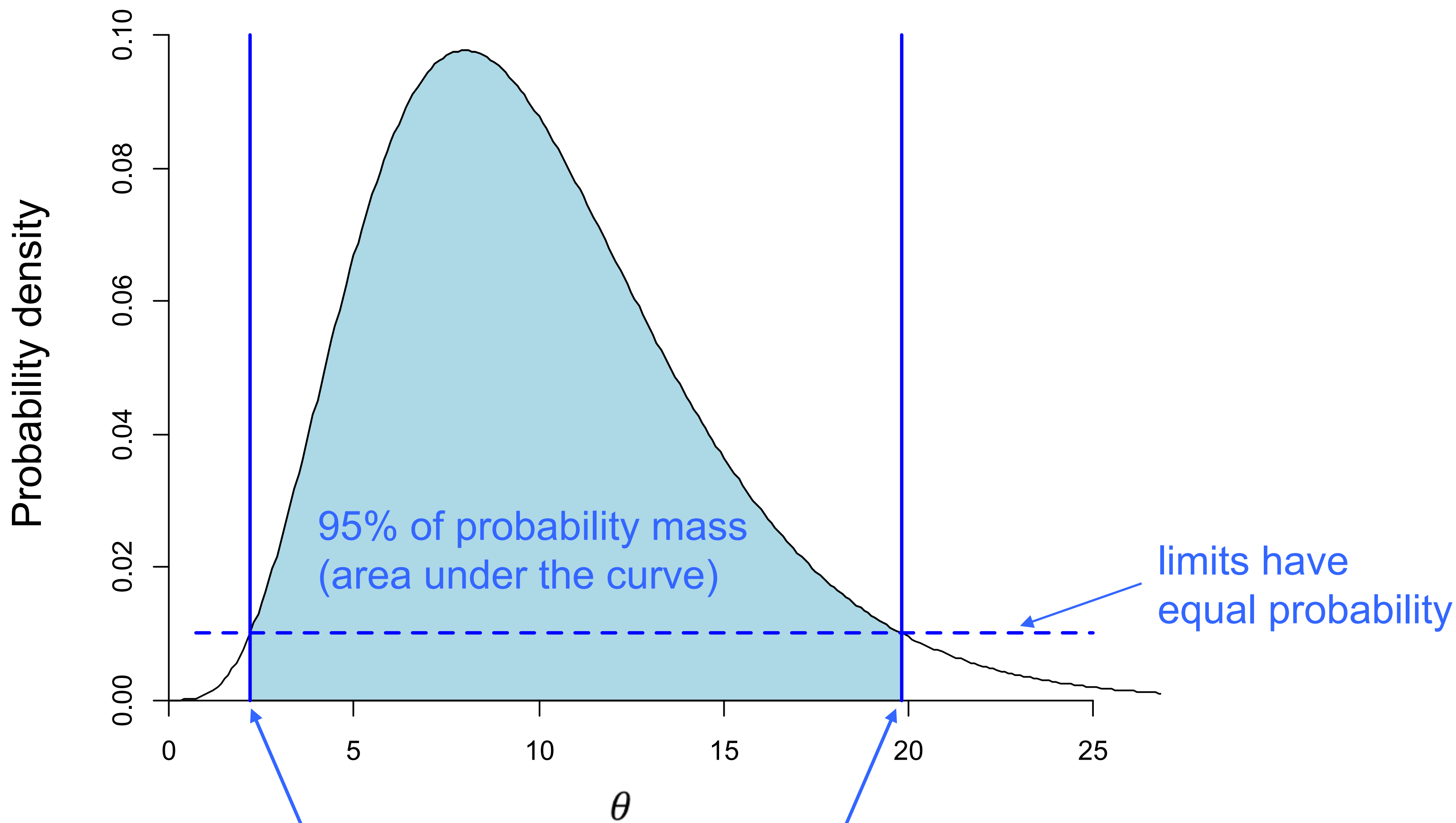
- Analogous idea to bootstrap samples but now we are sampling from the distribution of the parameter itself and not its sampling distribution

# Main messages from McElreath Ch 4

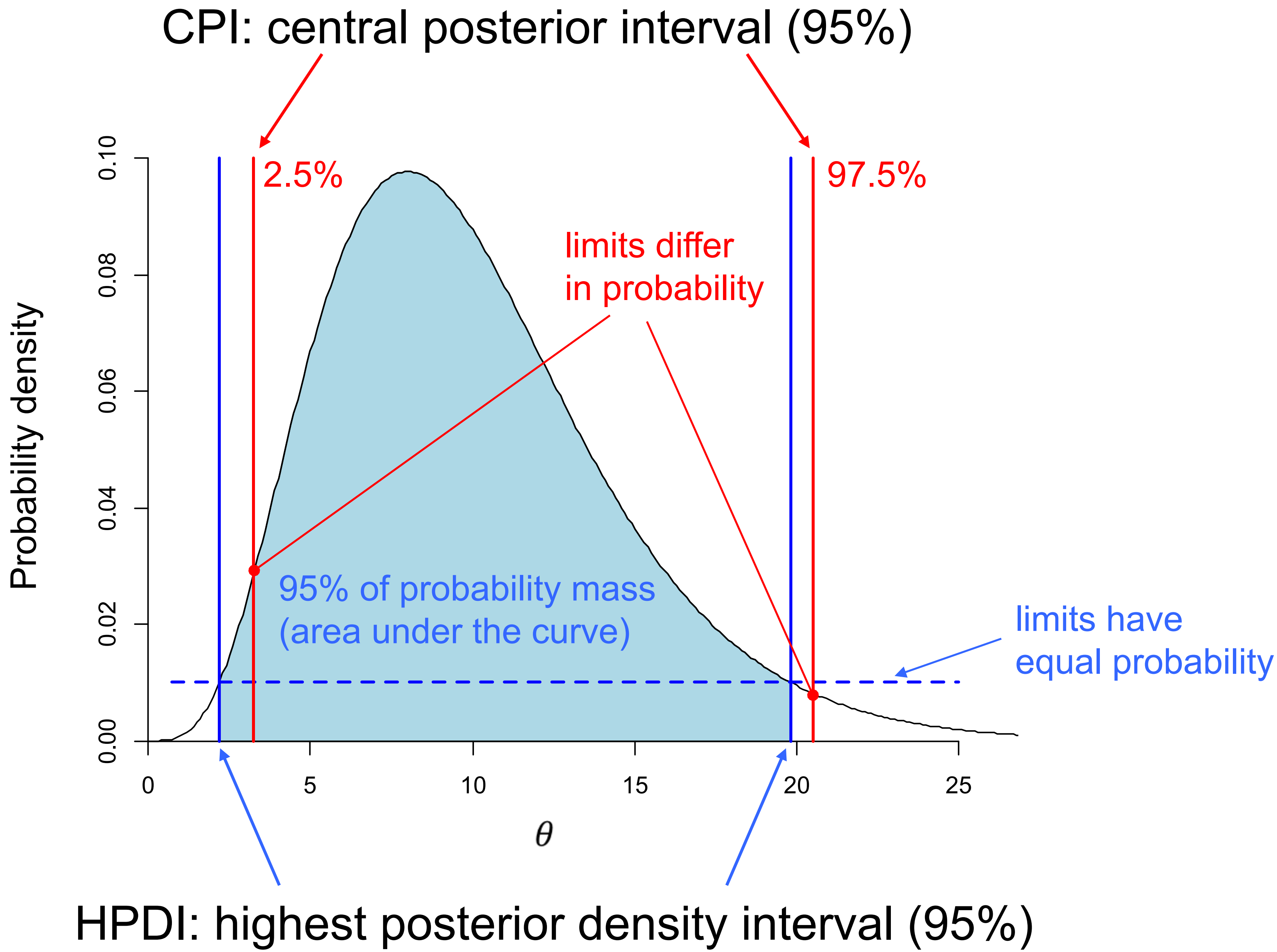
- Credible intervals are probabilities for model quantities (e.g. parameters or the average relationship)

# Credible intervals

- Plausibility intervals
- HPDI vs CPI




HPDI: highest posterior density interval (95%)


CPI: central posterior interval (95%)
2.5%
97.5%
limits differ
in probability
95% of probability mass
(area under the curve)
limits have
equal probability
Probability density
0.00
0.02
0.04
0.06
0.08
0.10
0
5
10
15
20
25
θ
HPDI: highest posterior density interval (95%)

# Main messages from McElreath Ch 4

- Credible intervals are probabilities for model quantities (e.g. parameters or the average relationship)
- Prediction intervals are probabilities for a new data point
  - Uncertainty in parameters + uncertainty in the data generating process

# Main messages from McElreath Ch 4

- You can derive quantities (e.g. height at weight 50 kg) from the posterior samples
  - any quantity that's a function of parameters
- Correlation among parameters can be obtained from the posterior samples
- Flat priors correspond to non-Bayesian approaches. Useful but there often better alternatives