BMI/CS 576 – Day 24

- Today
  - Bayesian network structure learning
    - Scoring networks
    - The space of possible networks
- Thursday
  - Sparse candidate algorithm

Course evaluation survey

- Please respond!
- Deadline: Dec 11th (Wed)
- Your feedback is critical to the future development of this course
- aefis.wisc.edu

Quiz

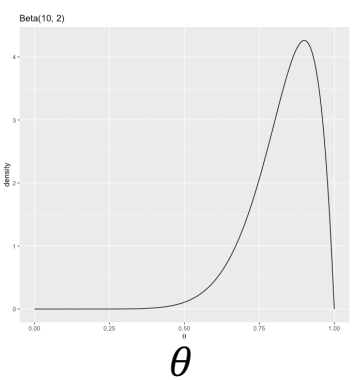
Suppose that you believe that a coin is likely to be weighted towards heads and that your prior belief in the probability of the coin coming up heads is $Beta(10, 2)$. You flip the coin k times, and, to your surprise, all flips come up tails. With a Bayesian perspective, what is the smallest value of k such that you would change your mind and believe that the coin is more likely weighted towards tails?

θ is the probability of a heads

$P(\theta) = Beta(10,2)$

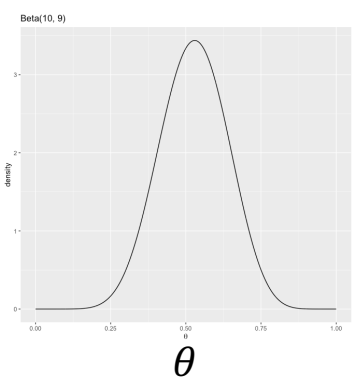
$P(\theta|D) = Beta(10, 2 + k)$

$k = 7$



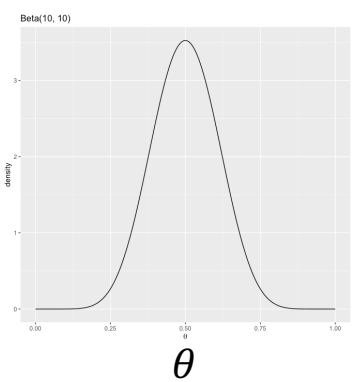
θ

$k = 8$

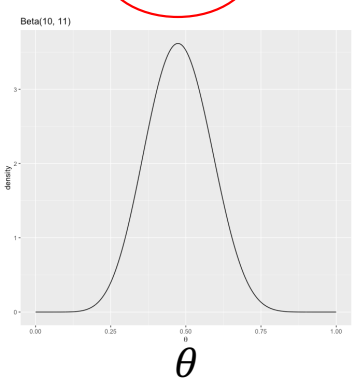


θ

$k = 9$



θ



θ

Questions – Model evidence

- P(D) or P(D|G)?
- When there is only one model, we will often just use P(D)
- When there are multiple possible models (represented by a Bayesian network graph structure G), we will often be explicit and say P(D|G)
  - Explicit about model evidence being a function of the model (G)
  - In this module we are optimizing over G

Conjugate priors

- What is a conjugate prior?
  - It is a form of the prior distribution such that the posterior distribution (after observing some data) has the same form
- Why is a conjugate prior useful?
  - Mostly for computational reasons
  - But also for ease in mathematical reasoning
  - Closed-form equations for many useful calculations
    - Posterior distribution
    - Model evidence