Overview of Bayesian inference

Stat 250

Click here for PDF version

A Bayesian "personality quiz"

Question 1

When flipping a fair coin, we say that "the probability of flipping Heads is 0.5." How do you interpret this probability?

1. If I flip this coin over and over, roughly 50% will be Heads.
2. Heads and Tails are equally plausible.
3. Both a and b make sense.

Question 2

An election is coming up and a pollster claims that "candidate A has a 0.9 probability of winning." How do you interpret this probability?

1. If we observe the election over and over, candidate A will win roughly 90% of the time.
2. Candidate A is much more likely to win than to lose.
3. The pollster's calculation is wrong. Candidate A will either win or lose, thus their probability of winning can only be 0 or 1.

Question 3

Consider two claims.

- Zuofu claims that he can predict the outcome of a coin flip. To test his claim, you flip a fair coin 10 times and he correctly predicts all 10.
- Kavya claims that she can distinguish natural and artificial sweeteners. To test her claim, you give her 10 sweetener samples and she correctly identifies each.

In light of these experiments, what do you conclude?

1. You're more confident in Kavya's claim than Zuofu's claim.
2. The evidence supporting Zuofu's claim is just as strong as the evidence supporting Kavya's claim.

Question 4

Suppose that during a recent doctor's visit, you tested positive for a very rare disease. If you only get to ask the doctor one question, which would it be?

1. What's the chance that I actually have the disease?
2. If in fact I don't have the disease, what's the chance that I would've gotten this positive test result?

Tally your points

Question 1:

- 1 = 1 points
- 2 = 3 points
- 3 = 2 points

Question 2:

- 1 = 1 points
- 2 = 3 points
- 3 = 1 points

Question 3:

- 1 = 3 points
- 2 = 1 points

Question 4:

- 1 = 3 points
- 2 = 1 points

What does your score mean?

- 4-5 → you're more of a frequentist thinker
- 6-8 → you see the merit in both (a pragmatist?)
- 9-12 → you're more of a Bayesian thinker

Question 1: Interpreting probability

When flipping a fair coin, we say that “the probability of flipping Heads is 0.5.” How do you interpret this probability?

1. (Frequentist) If I flip this coin over and over, roughly 50% will be Heads.
2. (Bayesian) Heads and Tails are equally plausible.
3. Both a and b make sense.

Question 2: Interpreting probability

An election is coming up and a pollster claims that "candidate A has a 0.9 probability of winning." How do you interpret this probability?

1. (Frequentist) If we observe the election over and over, candidate A will win roughly 90% of the time.
2. (Bayesian) Candidate A is much more likely to win than to lose.
3. (Rabid frequentist) The pollster's calculation is wrong. Candidate A will either win or lose, thus their probability of winning can only be 0 or 1.

Question 3: Balancing prior info and observed data

Consider two claims.

- Zuofu claims that he can predict the outcome of a coin flip. To test his claim, you flip a fair coin 10 times and he correctly predicts all 10.
- Kavya claims that she can distinguish natural and artificial sweeteners. To test her claim, you give her 10 sweetener samples and she correctly identifies each.

In light of these experiments, what do you conclude?

1. (Bayesian) You're more confident in Kavya's claim than Zuofu's claim.
2. (Frequentist) The evidence supporting Zuofu's claim is just as strong as the evidence supporting Kavya's claim.

Question 4: Asking questions

Suppose that during a recent doctor's visit, you tested positive for a very rare disease. If you only get to ask the doctor one question, which would it be?

1. (Bayesian) What's the chance that I actually have the disease?
2. (Frequentist) If in fact I don't have the disease, what's the chance that I would've gotten this positive test result?

A first Bayesian example

What proportion of Carleton students do you believe pulled at least one all-nighter during Fall term in order to get their school work done? Choose one of the options:

0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1

Overview of the Bayesian method

1. Choose (or elicit) a probability distribution to express the pre-data belief about the parameter of interest, θ.
2. Choose a model for the data given θ.
3. Observe data, $Y_1, \ldots, Y_n$.
4. Update the belief about θ by combining the prior belief and the data.
5. Draw inferences using this updated belief about θ.

What proportion of all Carleton students pulled at least one all-nighter to get school work done last term?

- Suppose we (incorrectly) assume that this class is representative of all Carleton students
- **Prior belief:** drawn from our survey
- **Data:** number of students in this class who pulled at least one all-nighter last term

Deriving the posterior

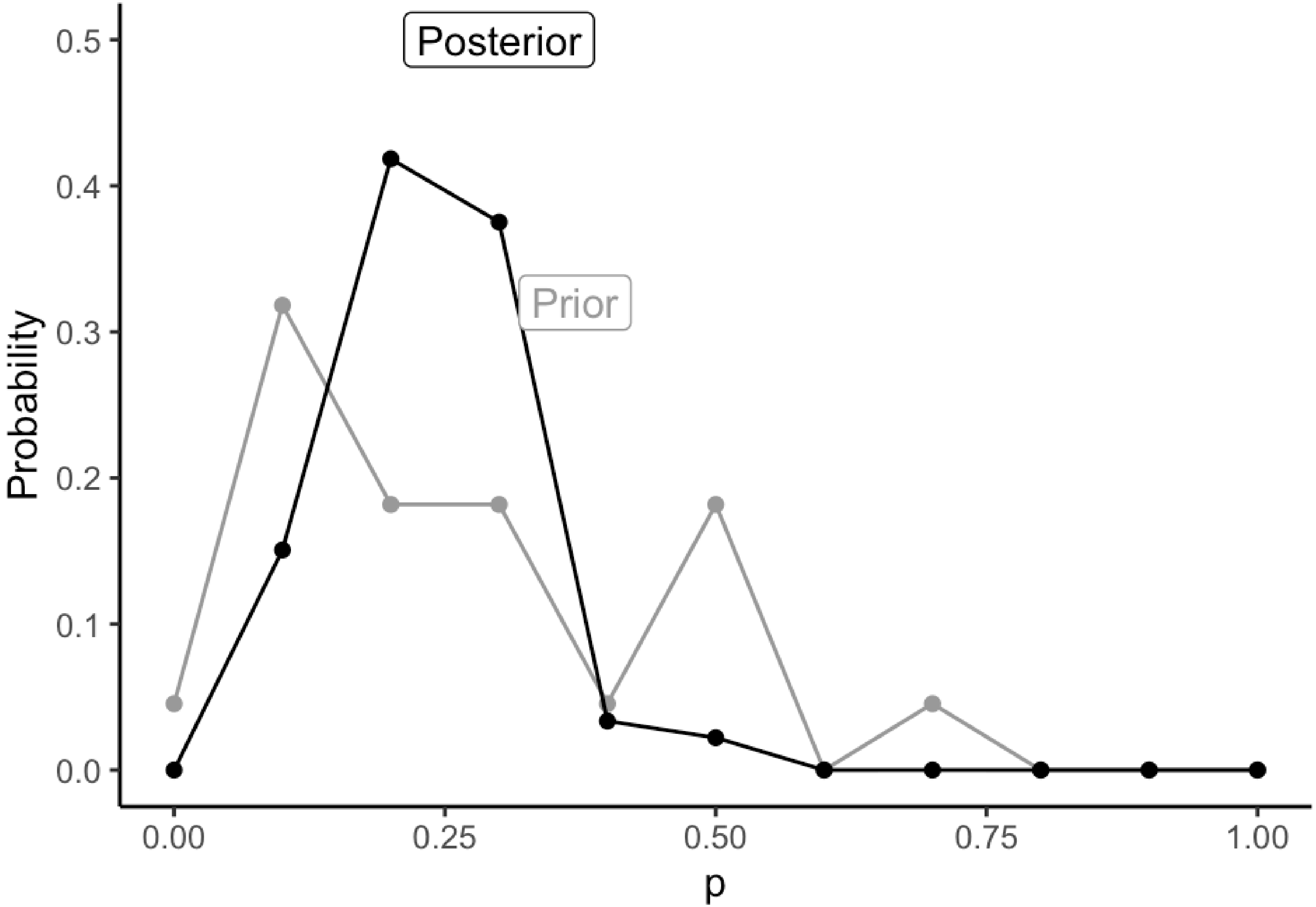
p	prior	likelihood	prior x likelihood	posterior
0.000000	0.045455	0.000000	0.000000	0.000000
0.100000	0.318182	0.037707	0.011998	0.150653
0.200000	0.181818	0.183287	0.033325	0.418457
0.300000	0.181818	0.164330	0.029878	0.375176
0.400000	0.045455	0.058785	0.002672	0.033552
0.500000	0.181818	0.009703	0.001764	0.022153
0.600000	0.000000	0.000680	0.000000	0.000000
0.700000	0.045455	0.000015	0.000001	0.000008
0.800000	0.000000	0.000000	0.000000	0.000000
0.900000	0.000000	0.000000	0.000000	0.000000

p	prior	likelihood	prior x likelihood	posterior
1.000000	0.000000	0.000000	0.000000	0.000000

Updating belief in R

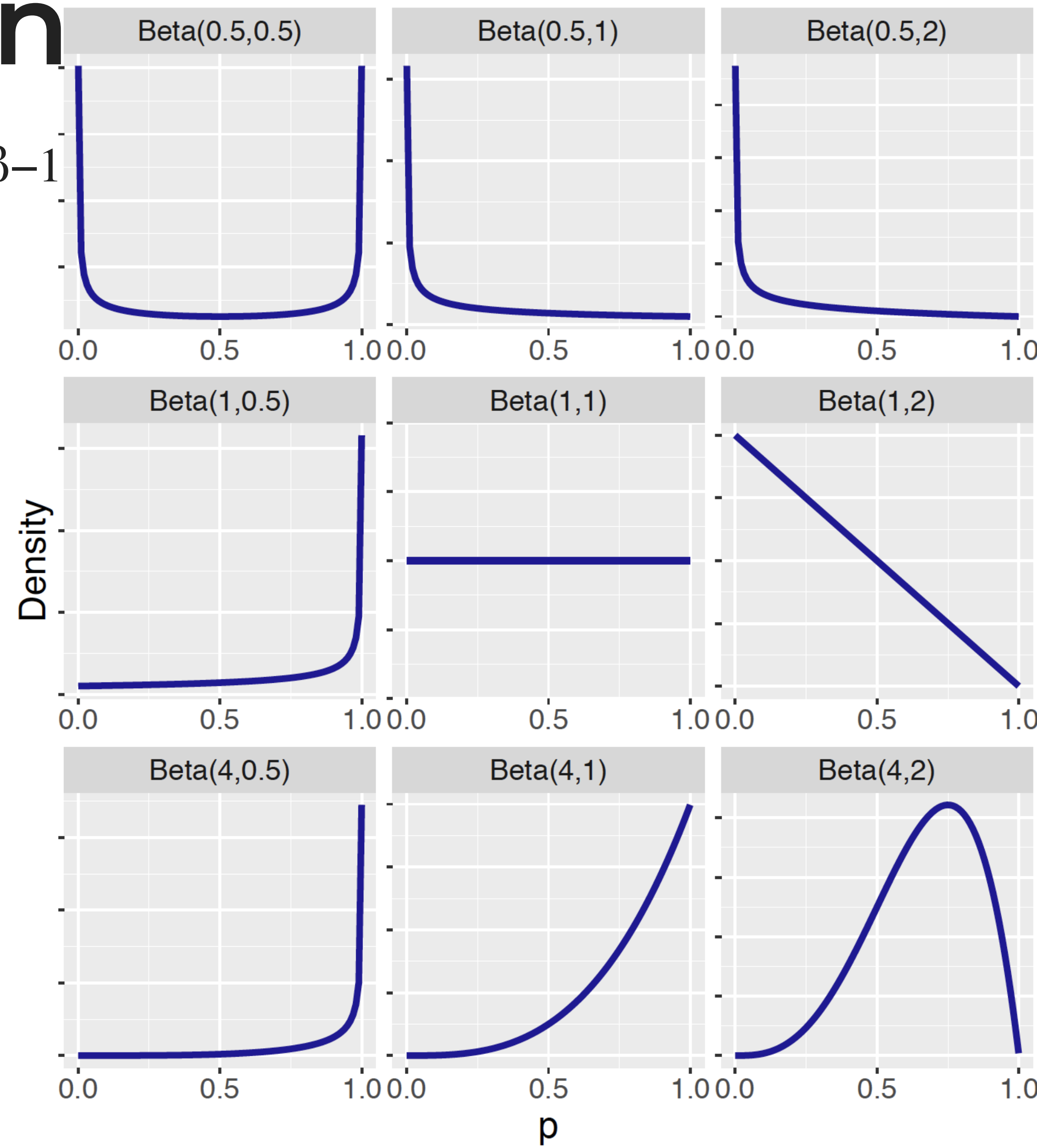
```r
# Prior based on class consensus
prior_values <- seq(0, 1, by = 0.1)
prior_probs  <- c(1, 7, 4, 4, 1, 4, 0, 1, 0, 0, 0) / 22

# Data
n <- 21 # sample size
y <- 5  # obs. no. of all-nighters

# Binomial likelihood
binom_lik <- dbinom(y, n, p = prior_values)

# Updating prior belief
joint_prob <- prior_probs * binom_lik  # prior x likelih
marginal_y <- sum(joint_prob)          # P(Y = y)
```

What have we accomplished?



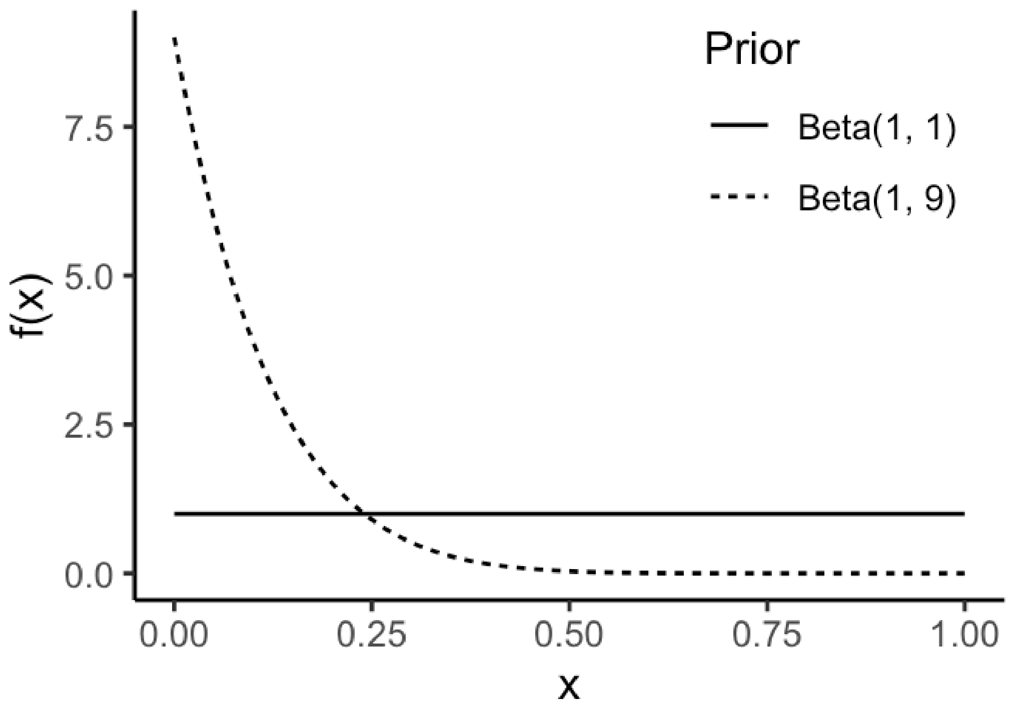
Beta distribution

- $f(x|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$
- Parameter space: $\alpha > 0, \beta > 0$
- Support: $0 < x < 1$

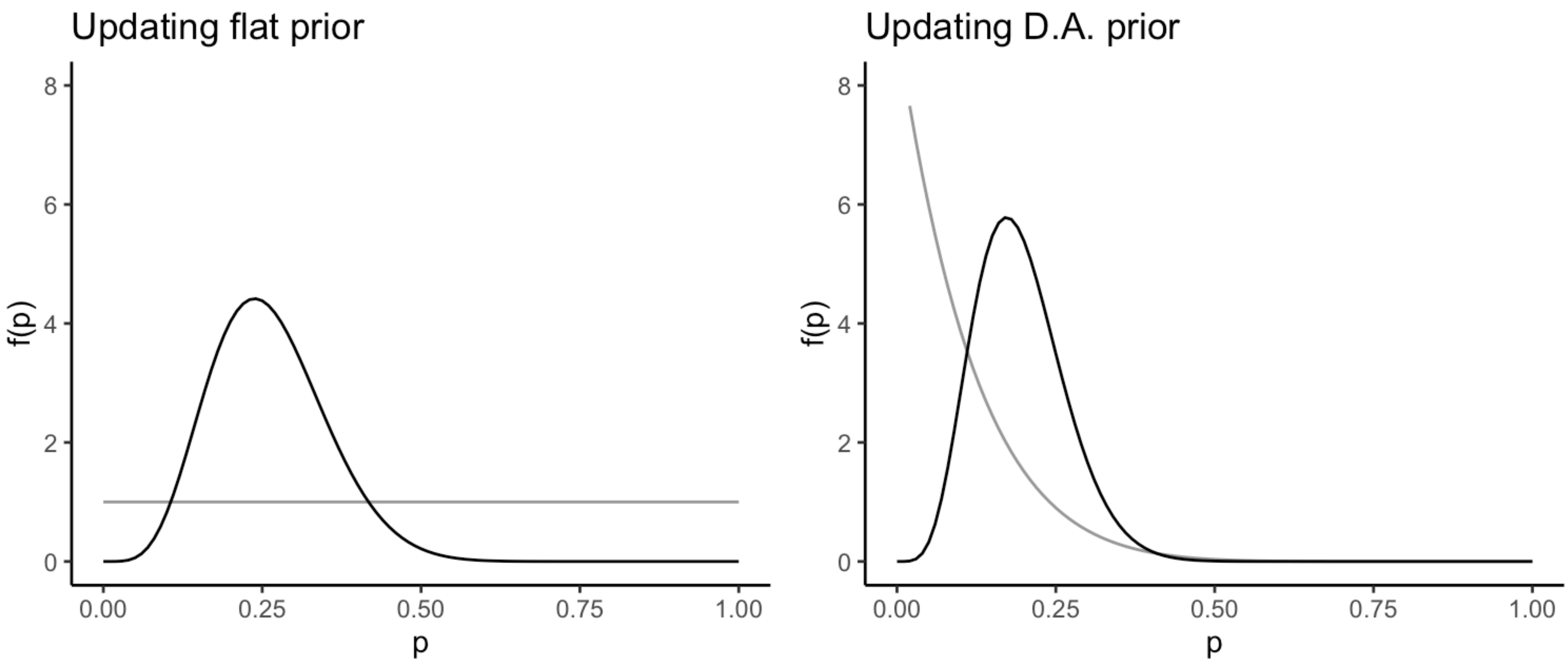


Beta priors

1. **Flat prior (non-informative)**: Beta(1,1)
2. **Data-augmentation prior**:
   Beta(# prior successes, # prior failures)



Posterior distributions



Credible intervals

In the Bayesian paradigm we treat the parameter as a random variable, so we can make probabilistic statements using the posterior distribution.

Using the flat prior...

```
qbeta(c(.025, .975), 1 + y, 1 + (n - y))
## [1] 0.1072892 0.4537036
```

Using the D.A. prior

```
qbeta(c(.025, .975), 1 + y, 9 + (n - y))
## [1] 0.07713551 0.34721170
```