GENOME 541 Section 2
Lecture 1
Introduction to Bayesian statistics

Yi Yin (yy224[at]uw.edu)
Shendure Lab
Department of Genome Sciences
University of Washington
Tuesday, April 11, 2017

Overview

- Based loosely on Peter Hoff's textbook: A First Course in Bayesian Statistical Methods
    1. Introduction to Bayesian statistics (Ch1 & 3)
    2. Gibbs sampling (Ch6)
    3. Metropolis-Hastings (Ch10)
    4. Bayesian linear models (Ch9)
- Most class notes will be presented on the board, take copious notes will help with HW.

Overview

- HW 3 (Apr. 12-Apr. 20) & HW 4 (Apr. 19-Apr. 27). LaTeX template will be provided but hand-written (<u>clear</u>) for derivation is OK. Generally, derive some posterior distribution, posterior predictive distribution, truncated distribution ... and simulate from it

- Familiarity of R is assumed (examples will be provided)

- Emphasis on <u>Bayesian inference</u> : is my coin fair? (example on board)

How do you go from $p(Y|\pi)$ to $p(\pi|Y)$?

Event A and B have probability $p(A)$ and $p(B)$

$$p(A, B) = p(A)p(B|A) = p(B)p(A|B)$$
$$p(\pi)p(Y|\pi) = p(Y)p(\pi|Y)$$

$$\boxed{p(\pi|Y) = \frac{p(\pi)p(Y|\pi)}{p(Y)}}$$

$p(\pi)$: prior probability of π
$p(Y|\pi)$: conditional probability of event/data given parameter π
$p(\pi|Y)$: posterior probability of π after observing some data Y
$p(Y)$: marginal probability of event/data Y

$$p(\pi|Y) = \frac{p(\pi)p(Y|\pi)}{p(Y)}$$

Provides a mechanism of learning from the data Y

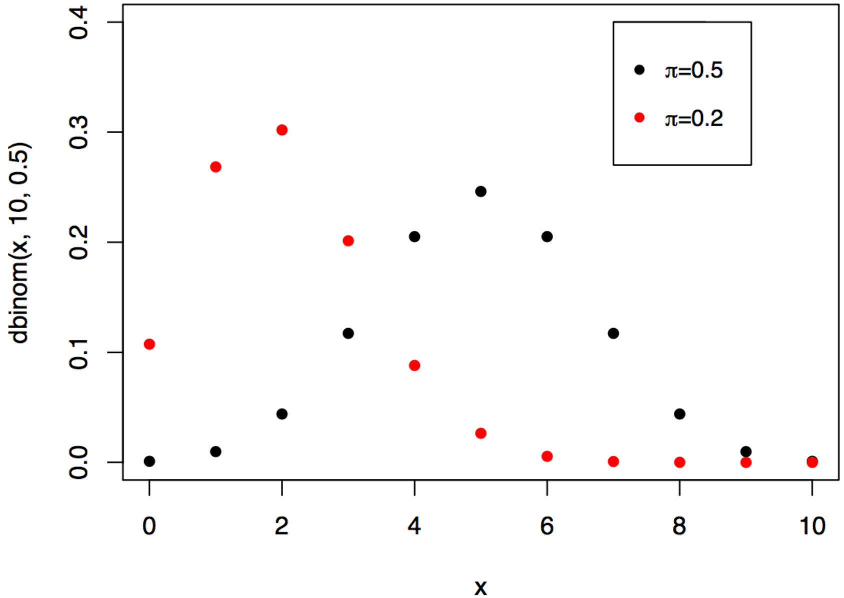
1. Prior distribution
   (for parameters in the model)

$$\xrightarrow[\text{Bayes' rule}]{\text{Apply}}$$ Posterior Distribution

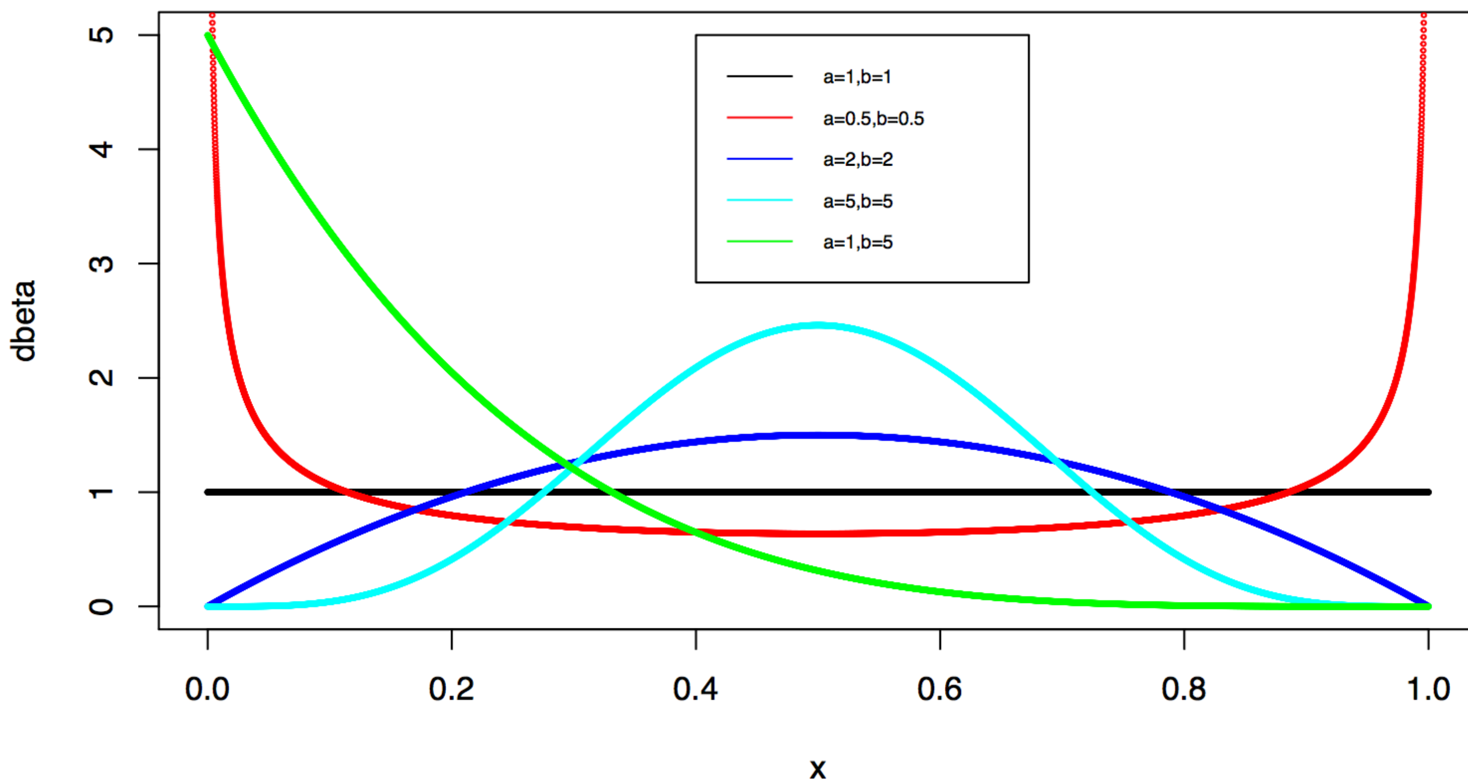
2. Likelihood function
   (for the data)

"conjugacy" in simple models: prior and posterior of the same form

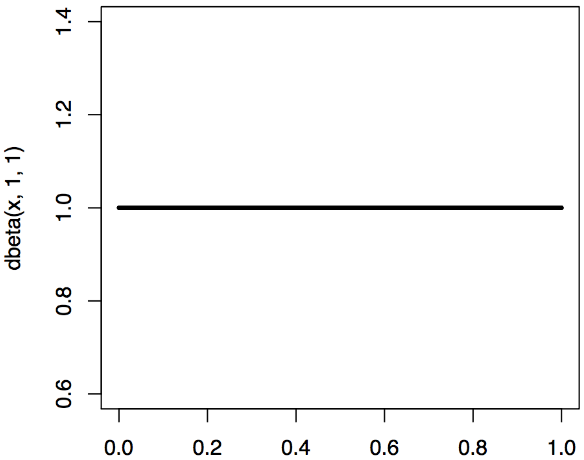
Most often involves unknown normalizing constant (is it necessary to calculate this? See example on board)



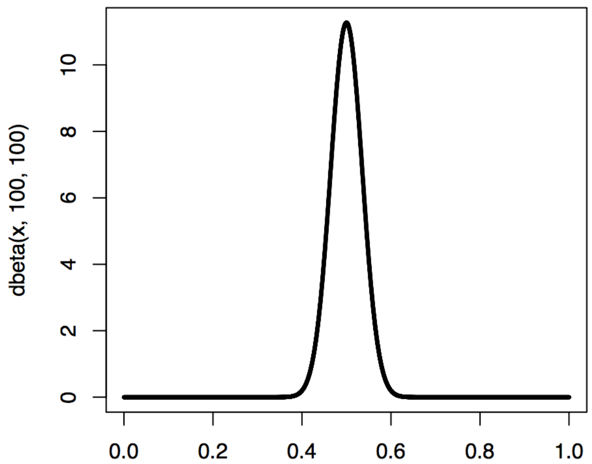
Binomial (10, π)



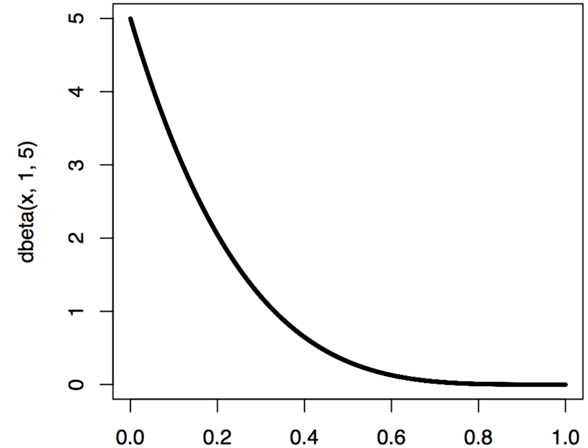
π ~ Beta(a,b)



$\pi_1 \sim \text{Beta}(1, 1)$
$E[\pi_1] = 0.5$

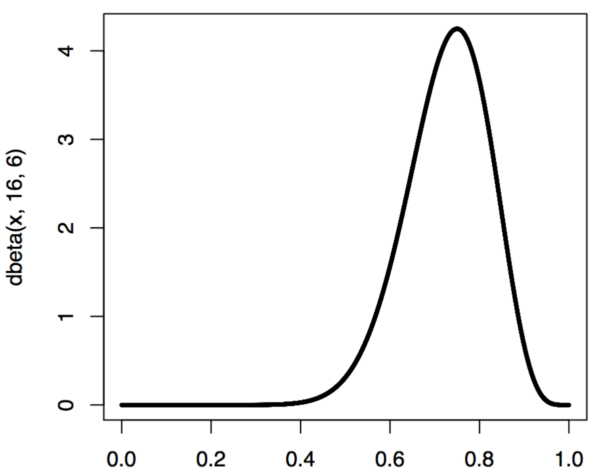


$\pi_2 \sim \text{Beta}(100,100)$
$E[\pi_2] = 0.5$

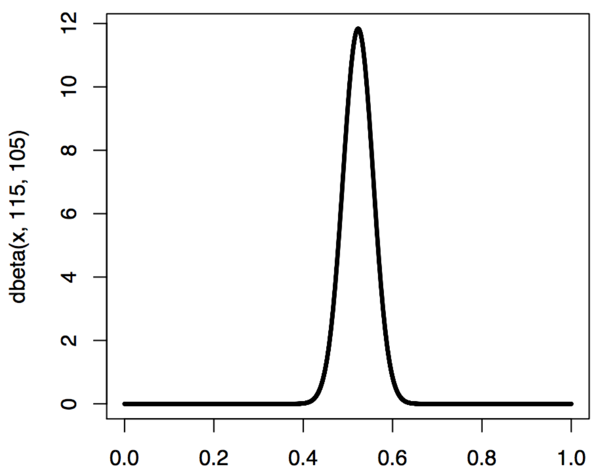


$\pi_3 \sim \text{Beta}(1,5)$
$E[\pi_3] = 0.167$

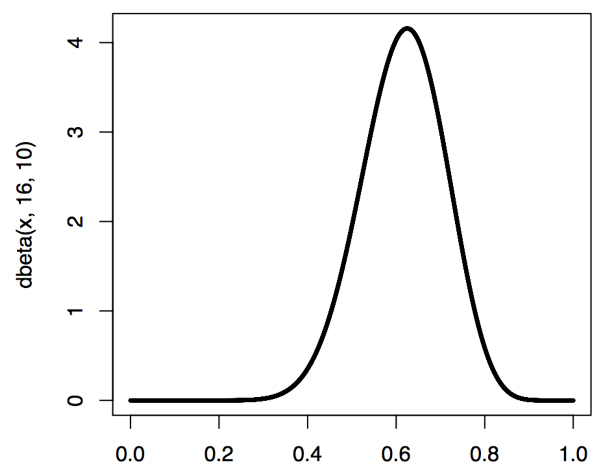
Now, we did some experiment and observed some data:
n=20, k=15; (15 head, 5 tail)



$\pi_1 | Y \sim \text{Beta}(16, 6)$
$E[\pi_1 | Y] = 0.73$



$\pi_2 | Y \sim \text{Beta}(115,105)$
$E[\pi_2 | Y] = 0.52$



$\pi_3 | Y \sim \text{Beta}(16,10)$
$E[\pi_3 | Y] = 0.615$

Let's look at the posterior mean:

Under π_1 ~ Beta (1, 1), after observing k=15, (n-k)=5, we get:
$\pi_1|Y$ ~ Beta (16, 6), with $E[\pi_1|Y]=16/(16+6)$.

Under π_1 ~ Beta (a, b), after observing k heads, (n-k) tails, we get:
$\pi_1|Y$ ~ Beta (a+k, b+(n-k)), with $E[\pi_1|Y]=(a+k)/(a+k+b+(n-k))=(a+k)/(a+b+n)$

$$E[\pi_1|Y] = \frac{a+k}{a+b+n} = \frac{a}{a+b} \times \frac{a+b}{a+b+n} + \frac{k}{n} \times \frac{n}{a+b+n}$$

a+b: prior sample size (2, 100 and 6 in the last example)
n: sample size in data

Posterior mean is the weighted average of prior mean and sample mean.

1. Prior specification (will cover examples on the impact of eliciting different priors)(Lindley paradox…)

2. Choice of likelihood function (same as in frequentist analyses, not covered here, but will usually be provided in HWs)

3. Derive and/or simulate from posterior (focus) (Posterior needs to be a **<u>proper</u>** distribution!!) (HW2)

4. Inferences (based on posterior and loss function)

1. Any questions?

2. HW1: Poisson distribution with Gamma prior (understand the relationship between Poisson distribution and Negative Binomial distribution)

3. HW1 (ungraded question): think about and describe in words your ideas on how to sample from a truncated distribution