Example Bayesian Model:

$$y_i \sim \text{Binom}(N_i, p), \quad i = 1, \ldots, n \quad \text{(groups)}$$

$$p \sim \text{Beta}(\alpha, \beta)$$

# of surviving individuals

# of initial individuals

survival probability

Posterior:

$$[p \mid y] = \frac{[y \mid p][p]}{[y]}$$

$$\propto \prod_{i=1}^{n} [y_i \mid p][p]$$

$$\propto \left(\prod_{i=1}^{n} \cancel{\binom{N_i}{y_i}} p^{y_i} (1-p)^{N_i - y_i} \right) \cancel{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}} p^{\alpha-1} (1-p)^{\beta-1}$$

$$\propto p^{\sum_i y_i} (1-p)^{\sum_i N_i - y_i} \; p^{\alpha-1} (1-p)^{\beta-1}$$

$$\propto p^{\sum_i y_i + \alpha - 1} (1-p)^{\sum_i (N_i - y_i) + \beta - 1}$$

$$= \text{Beta}\left(\sum_{i=1}^{n} y_i + \alpha, \sum_{i=1}^{n} (N_i - y_i) + \beta \right)$$

Because we recognize the posterior distribution as a known form and hence know its normalizing constant!

Introduction:

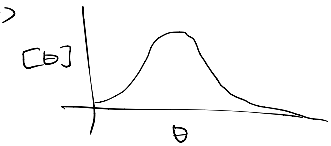
Bayesian Basics:

Probability distributions characterize stochasticity of random variables: $\theta \sim [\theta]$

pdf or pmf

<u>Note:</u> $[\theta] = f(\theta) = P(\theta) = \pi(\theta) \Rightarrow$

<u>and</u> $\int [\theta] d\theta = 1$



When 2 quantities we can find conditional distribution of one given the other:

$$[\theta | y] = \frac{[y|\theta][\theta]}{[y]}$$

data model

parameter model

posterior

$$\propto [y|\theta][\theta]$$

proportional w.r.t. θ

<u>Note:</u> $[y] = \int [y, \theta] d\theta = \int [y|\theta][\theta] d\theta$

marginal distn of data (also prior predictive distn. and "evidence")

Plug marginal into cond. prob to get Bayes rule:

$$[\theta | y] = \frac{[y|\theta][\theta]}{\int [y|\theta][\theta] d\theta},$$

we usually only deal w/ numerator b/c denom. is hard to calculate.

<u>Bayesian model:</u> $\quad y \sim [y|\theta]$ data model "likelihood"

$\theta \sim [\theta]$ prior

Conditioning makes it easy to specify! usually $(y,\theta) \sim [y,\theta]$ is hard to specify directly.

The Bayesian model is generative! Can simulate data y from it.

<u>First</u>: sample θ from $[\theta]$

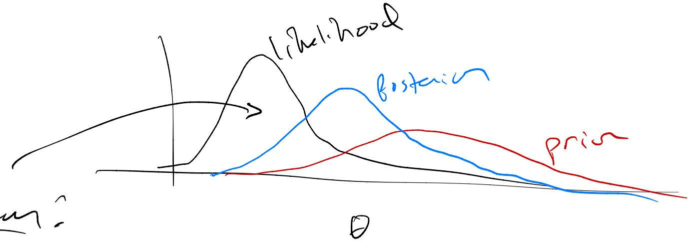
<u>Second</u>: sample y from $[y|\theta]$

Probability distributions are essential elements of Bayesian models. Lots to choose from, so memorize common ones: binomial, Poisson, Normal, Gamma, inv. gamma, multivariate Normal, etc...

(you can also invent prob. distributions if you need something special).

<u>Key distinction</u>: θ and y are <u>random</u> before they are observed. We never observe θ, so we use posterior $[\theta|y]$ to learn about it.

Posterior distn is optimal blend of likelihood and prior.



problem:

$\int [y|\theta] d\theta \neq 1$