Bayes' theorem tells us how to update our beliefs (about unknown parameters, θ) w/ data.

in this class

$$p(\theta) \longrightarrow p(\theta \mid y)$$

data

unknown parameter of the data generative model.

$$p(\theta \mid y) = \frac{\boxed{p(y \mid \theta)}\, p(\theta)}{\int p(y \mid \theta)\, p(\theta)\, d\theta}$$

likelihood

prior

posterior

constant: $\int p(y, \theta)\, d\theta = p(y)$

not a function of θ!

"marginal likelihood"

Typical scenario for a (Bayesian) statistician:

here's a data set: $0, 1, 0, 0, \ldots$

Ask yourself: what simple model could have generated this data

obvious answer here: "binary" model: $x_i \mid \theta \overset{iid}{\sim} \text{binary}(\theta)$

i.e. $p(x_i = 1) = \theta$
$p(x_i = 0) = (1 - \theta)$ and $\theta \in [0, 1]$

$$\Rightarrow p(x_1, \ldots, x_n \mid \theta) = \theta^{\sum x_i} (1 - \theta)^{n - \sum x_i}$$

- always start w/ simplest model
- when in doubt, go to distr. sheet & look at support of various distributions.
  ... like they could have generated your data?

2) Identify what you don't know in the likelihood.

Here : $\boxed{\theta}$

<u>and</u> write down a prior for all unknowns

(consider the support of θ).

Candidate priors :

- Unif $(0,1)$
- beta(a,b)
- truncated normal

---

we'll begin by considering the uniform:

$$p(\theta) = \begin{cases} 1 & \text{if } \theta \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

Together with our likelihood:

$$p(y|\theta) = \binom{n}{y} \theta^y (1-\theta)^{n-y}$$

← re-written as binomial according to transform $y = \sum_{i=1}^{n} x_i$

we can directly write down our posterior (it is deterministically defined by likelihood & prior)

$$p(\theta|y) = \frac{\binom{n}{y} \theta^y (1-\theta)^{n-y} \cdot \mathbb{1}_{\{\theta \in [0,1]\}}}{p(y)}$$

But notice :

$$\boxed{p(\theta|y) \propto \theta^y (1-\theta)^{n-y} \qquad \theta \in [0,1]}$$

proportional in θ!

This is the kernel of a <u>beta(α,β)</u> $\quad \alpha = y+1$

$\beta = n-y+1$

Prove claim: beta prior is conjugate to binomial data generative model.

In other words, show that

$\theta \sim \text{beta}(a,b)$

$y|\theta \sim \text{binomial}(n,\theta)$

implies $\theta|y \sim \text{beta}(\alpha,\beta)$ for some α,β.

<u>Note</u>: if we show that the posterior is proportional to the <u>kernel</u> of a beta density, we will be done.

<u>Proof</u>

$$p(\theta|y) \propto \text{likelihood} \cdot \text{prior}$$
$$\propto \theta^{y}(1-\theta)^{n-y} \cdot \theta^{a-1}(1-\theta)^{b-1}$$
$$\propto \theta^{y+a-1}(1-\theta)^{(n-y+b)-1}$$

$$\Rightarrow \theta|y \sim \text{beta}(y+a,\, n-y+b)$$

$\square$

<u>Exercise</u>

show that for $\theta|y \sim \text{beta}(y+a, n-y+b)$

$$\lim_{n\to\infty} \mathbb{E}[\theta|y] = \bar{x} \quad \text{where } \bar{x} = \frac{1}{n}\sum x_i$$

<u>roof</u>

$$\mathbb{E}[\theta|y] = \frac{a+y}{a+b+n}, \text{ notice } y = n\bar{x},$$

$$\lim_{n\to\infty} \frac{a+n\bar{x}}{a+b+n} = \lim_{n\to\infty} \underbrace{\frac{a}{a+b+n}}_{0} + \frac{n\bar{x}}{a+b+n}\cdot\frac{1/n}{1/n} = \frac{\bar{x}}{1} = \bar{x}$$

$\square$