Bayes' thm tells us how to

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

$y = \sum x_i$

update beliefs about θ w/ data.

$\underline{\underline{\theta}}$ is called a "parameter" of the data generative model.

$$\underbrace{p(\theta \mid y)}_{\boxed{\text{posterior}}} = \frac{\overbrace{\boxed{p(y \mid \theta)}}^{\boxed{\text{likelihood}}}\ \overbrace{p(\theta)}^{\boxed{\text{prior}}}}{\underbrace{\int p(y \mid \theta)\, p(\theta)\, d\theta}_{\text{"constant"}}}$$

$$\underbrace{\int p(y, \theta)\, d\theta = p(y)}_{\text{Not a function of } \theta!}$$

What's a suitable prior for θ?

$\theta \in (0, 1)$.

beta distr. or uniform $(0, 1)$ are both good candidates.

D.G.M: $y \sim \text{binomial}(n, \theta)$

complete Bayesian model

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

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

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

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

This is the kernel of a $\underline{\text{beta}(\alpha, \beta)}$

$\alpha = y+1 \quad \beta = n-y+1$

Prove claim: beta prior is conjugate to binomial data generative model. $\theta \sim \text{beta}(a, b)$

want to show: posterior has a beta kernel.

$$p(\theta|y) \propto \theta^y (1-\theta)^{n-y} \theta^{a-1} (1-\theta)^{b-1}$$

$$\propto \theta^{y+a-1} (1-\theta)^{(n-y+b)-1}$$

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

$y = \Sigma x_i$ $\qquad \bar{x} = \frac{1}{n} \Sigma x_i$

$= n\bar{x}$

$$\lim_{n\to\infty} \mathbb{E}\,\theta | y = \frac{a+y}{a+b+n}$$

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

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

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

$$= \bar{x}$$

EX: write posterior mean as

$w\,\mathbb{E}(\theta) + (1-w)\bar{x}$

$$1-w = \frac{n}{a+b+n} \Rightarrow w = 1-(1-w)$$

$$= \frac{a+b+n}{a+b+n} - \frac{n}{a+b+n}$$

$$= \frac{a+b}{a+b+n}$$

$$\underbrace{\left(\frac{a+b}{a+b+n}\right)}_{w} \cdot \underbrace{\frac{a}{a+b}}_{\mathbb{E}(\theta)} + \underbrace{\frac{n}{a+b+n}}_{(1-w)} \cdot \bar{x}$$