$$\underbrace{p(\theta \mid x)}_{\substack{\text{posterior} \\ \text{dist}^{\underline{n}}}} = \frac{p(\theta, x)}{\underbrace{p(x)}_{\substack{\text{marginal} \\ \text{likelihood} \\ \text{of the data} \\ x}}} = \frac{\underbrace{p(x \mid \theta)}_{\text{likelihood}} \underbrace{p(\theta)}_{\text{prior}}}{\underbrace{\boxed{p(x)}}_{\text{"marginal"}}}$$

fixed

θ: random variable

→ dropping $p(x)$ since it's constant or fixed wrt to θ

$$\underset{\theta}{\propto} \underbrace{p(x \mid \theta)}_{\text{likelihood}} \underbrace{p(\theta)}_{\text{prior}}$$

Example:

Bernoulli distribution: commonly used for binary outcomes such as flipping a coin.

$$X \sim \text{Bernoulli}(\theta), \quad 0 < \theta < 1$$

$$\underline{p(x \mid \theta)} = \underline{\theta^{x} (1-\theta)^{1-x}}, \quad 0 < \theta < 1.$$

prob of success (heads)

n = # of obs.

$$X_1, \dots, X_n \overset{\text{iid}}{\sim} \text{Bernoulli}(\theta)$$

$$p(x_1, \dots, x_n \mid \theta) = \prod_{i=1}^{n} \underline{P(X_i = x_i \mid \theta)}$$

$$= \prod_{i=1}^{n} p(x_i \mid \theta)$$

$1-x_i$

$$= \prod_{i=1}^{n} \left[\theta^{x_i} (1-\theta) \right]$$ ← how can I simplify

$$\boxed{p(x_{1:n} \mid \theta) = \theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n - \sum_{i=1}^{n} x_i} \quad \text{(likelihood)}} \quad (1)$$

Instead of performing frequentist inference which assume that θ is fixed but unknown we're going to assume θ is random + unknown. ⇒ we're going to put a dist on θ!

$$\theta \sim \text{Beta}(a, b), \quad a, b > 0 \text{ known}$$

$$\boxed{p(\theta) := \text{Beta}(\theta \mid a, b) = \frac{1}{B(a,b)} \theta^{a-1} (1-\theta)^{b-1}} \quad (2)$$

$0 < \theta < 1$

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$X_1, \ldots, X_n \mid \theta \overset{iid}{\sim} \text{Bernoulli}(\theta)$$

$$\theta \sim \text{Beta}(a, b), \quad \underset{\text{known}}{a, b > 0}$$

Goal: Find $p(\theta \mid x_{1:n})$ (posterior distribution)

Soln:

$$p(\theta \mid x_{1:n}) \underset{\theta}{\propto} \underbrace{p(x_{1:n} \mid \theta)}_{\text{likelihood}} \underbrace{p(\theta)}_{\text{prior}}$$

Bayes' Thm

$$\underset{\theta}{\propto} \underbrace{\theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i}}_{\text{likelihood}} \times \underbrace{\frac{1}{B(a,b)} \theta^{a-1} (1-\theta)^{b-1}}_{\text{prior}}$$

$$\underset{\theta}{\propto} \theta^{\sum x_i + a - 1} (1-\theta)^{n - \sum x_i + b - 1} \times \boxed{\frac{1}{B(a,b)}}$$

constant wrt θ (drop it)

$$\underset{\theta}{\propto} \theta^{\sum x_i + a - 1} (1-\theta)^{n - \sum x_i + b - 1}$$

(try + recognize this as a distn I know)

$\theta^{a-1} (1-\theta)^{b-1}$ updated $\text{Beta}(\theta \mid \sum x_i + a, n - \sum x_i + b)$

$$\Rightarrow \theta \mid x_{1:n} \sim \text{Beta}(\sum x_i + a, n - \sum x_i + b).$$

Why is like = dbeta (th, x+1, n-x+1)?

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

goal: write as a fn of θ so I can plot likelih, prior, & post together!

$$\underset{\theta}{\propto} \theta^{\underline{x+1}-1} (1-\theta)^{\underline{n-x+1}-1}$$

$\rightarrow$ updated beta dist (x+1, n-x+1)