Binomial Distribution:

$X \sim \text{Binomial}(n, p)$

X is a random variable

$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$

$\sum_{x=0}^{n} f_X(x) = 1$

Use θ instead of p and Θ (capital theta) for P throughout

Beta Distribution:

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

p is a random variable

$$f_p(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{(\alpha-1)} (1-p)^{(\beta-1)} = c \cdot p^{(\alpha-1)} (1-p)^{(\beta-1)}$$

whatever constant is necessary so that the density integrates to 1

on the support $0 \le p \le 1$

$\int_0^1 f_p(p)\, dp = 1$

~~Suppose we adopt the prior distribution~~

Let X be the # of blue M&M's in a sample of size n.

$X|p \sim \text{Binomial}(n, p)$

p is an unknown parameter we would like to estimate.

We express the state of our knowledge about p before observing any data in the prior distribution

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

(we pick α and β to express our beliefs)

Suppose we take a sample and observe x blue M&M's. How should we update our beliefs?

Posterior distribution of p:

$$f_{P|X}(p|x) = \frac{f_{P,X}(p,x)}{f_X(x)} = \frac{f_{P,X}(p,x)}{\int_0 f_{P,X}(p,x)\,dp}$$

x is observed, P is the r.v.

$$= \frac{f_P(p)\,f_{X|P}(x|p)}{\int_0 f_P(p)\cdot f_{X|P}(x|p)\,dp}$$

$$= \frac{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, p^{(\alpha-1)}(1-p)^{(\beta-1)}\cdot\binom{n}{x}p^{x}(1-p)^{n-x}}{\int_0 f_P(p)\cdot f_{X|P}(x|p)\,dp}$$

includes all terms in previous line that don't involve p including denominator

$$= C\; p^{(\alpha-1)}(1-p)^{(\beta-1)}\cdot p^{x}(1-p)^{n-x}$$

$$= C\; p^{(\alpha+x-1)}(1-p)^{(\beta+n-x-1)} \qquad (1)$$

Compare to the pdf of a Beta$(\alpha^{posterior}, \beta^{posterior})$ distribution:

$$f_P(p) = \frac{\Gamma(\alpha^{post}+\beta^{post})}{\Gamma(\alpha^{post})\Gamma(\beta^{post})}\, p^{(\alpha^{post}-1)}(1-p)^{(\beta^{post}-1)} \qquad (2)$$

Equations (1) and (2) match if

$$\alpha^{post} = \alpha + x \quad \text{and} \quad \beta^{post} = \beta + n - x$$

The <u>posterior</u> <u>distribution</u> for P, if a Beta prior is used, is

$$\Theta \mid X = x \sim \text{Beta}(\alpha+x,\ \beta+n-x)$$

<u>Def.:</u> If the posterior distribution and the prior distribution are in the same parametric family, the prior was a <u>conjugate prior</u> for the data model.

Example: The Beta distribution is a conjugate prior for a Binomial model.

↳ When we used a Beta prior, the posterior was also a Beta distribution.